

# THIS WEEK

## EDITORIALS

**LITERATURE** Thomas Pynchon splendidly unites the two cultures **p.282**

**WORLD VIEW** Beware the precarious state of the Himalayas **p.283**



**PREEN SPIRIT** Feather oil signature can indicate bird fertility **p.285**

## The final assessment

*The Intergovernmental Panel on Climate Change has provided invaluable evidence for policy-makers, but giant reports should give way to nimbler, more relevant research.*

**T**he first working group of the Intergovernmental Panel on Climate Change (IPCC) will deliver its assessment of the science of global warming at a meeting in Stockholm next week. This will be the fifth time that the IPCC has delivered such an assessment; some 23 years have passed since the first effort. Many things have changed in that time; others have not. Regardless, it is time to rethink the IPCC. The organization deserves thanks and respect from all who care about the principle of evidence-based policy-making, but the current report should be its last mega-assessment.

For more than two decades, the depth and breadth of the IPCC's regular reports have expanded exponentially and in parallel with a truly breathtaking array of data. More climate models are running increasingly sophisticated calculations, and coordinated experiments are bolstering our understanding of the results. Most importantly, the panel has increased its confidence in the underlying message — that greenhouse gases are altering Earth's climate. No serious politician on the planet can now dispute that.

Unfortunately, one thing that has not changed is that scientists cannot say with any certainty what rate of warming might be expected, or what effects humanity might want to prepare for, hedge against or avoid at all costs. In particular, the temperature range of the warming that would result from a doubling of atmospheric carbon dioxide levels is expected to be judged as 1.5–4.5°C in next week's report — wider than in the last assessment and exactly what it was in the report of 1990. The governments of the world, to whom the IPCC reports, have made precious little headway in reducing emissions. And they appear in little hurry to do so. For all of these reasons, it would seem that a little reform is in order.

After the first working group publishes its findings next week, attention will turn to the second and third groups, which focus on impacts and mitigation, respectively, and are scheduled to deliver their results next year. The result of this process will be a kind of consensus document that scientists, non-governmental organizations, bureaucrats and elected officials around the world can turn to as they discuss — at times it can seem endlessly — how to confront the most complex environmental issue of our time.

### OLD NEWS

The IPCC process remains a human endeavour and, as such, is subject to human error; the silly mistake in the previous report that Himalayan glaciers would melt completely by 2035 demonstrates this. But the rarity of such errors shows what a solid job the organization has done. Critics went through the rest of the more-than-900-page report with a fine-tooth comb but found little else of significance to crow about.

True, 'consensus' does not necessarily mean that everybody is entirely happy with judgments about how the science is framed. Many researchers felt that the fourth assessment underplayed the potential for rapid sea-level

rise, for instance, and this time around, some fear that the IPCC is putting too much weight on a series of studies suggesting that the climate may be less sensitive to greenhouse gases than previously thought. In the end, however, it is abundantly clear that the IPCC has done its job and is delivering what international policy-makers need to do theirs. Yes, greenhouse gases are changing the climate. Yes, we are already seeing substantial impacts, and more are on the way. And

yes, this adds up to a problem for society that is significant and warrants immediate attention.

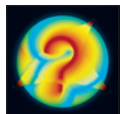
***"The IPCC's mega-assessments are out of date by the time they hit the streets."***

But none of this is news, and that is the problem. The IPCC's fifth assessment will provide a comprehensive analysis of policy options and the scientific basis for the next round of climate negotiations, which are scheduled to come to a head in 2015. What is missing from these talks is not science but political ambition, which is ultimately a reflection of public support. The IPCC has a crucial role in this process and must remain the central authority on global warming. It is not clear, however, that to immediately launch into yet another comprehensive assessment — which would consume immeasurable time and energy, and would probably come to the same bottom-line conclusions — represents the best use of our scientific resources.

Instead, climate scientists should focus on smaller and more rapid assessments of more pressing questions that have a particular political interest and for which science is evolving quickly. These reports could look more like the panel's recent special report on extreme weather; longer and more detailed assessments could be performed as needed, when there is sufficient interest from the governments that the IPCC serves.

Such a structure might also help to avoid an unfortunate consequence of the current framework, which ensures that the IPCC's mega-assessments are out of date by the time they hit the streets. For the latest document, some 20 international teams participated in coordinated modelling experiments, providing the core climate projections that the global community will use in the coming years; this is one area in which the IPCC has clearly driven the science forward. However, owing to logistics and deadlines, scientists barely had time to conduct a preliminary analysis for the current assessment, and as a result it lacks the more detailed analyses and most of the new science being published in journals today.

Absent from next week's report, for instance, is recent and ongoing research on the rate of warming and what is — or is not — behind the plateau in average global temperatures that the world has experienced during the past 15 years. These questions have important policy implications, and the IPCC is the right body to answer them. But it need not wait six years to do so. ■



**OUTLOOK FOR EARTH**

A Nature special issue on the IPCC  
[nature.com/ipcc2013](http://nature.com/ipcc2013)

# Spooked

Researchers and lawmakers must work to rebuild trust in secure Internet standards.

When John Hopkins University ordered cryptography researcher Matthew Green to take down a blog post last week, it found that its action only made the material more visible. The university quickly backed down, but the global media began to pay attention to the post, which discussed revelations that the US National Security Agency (NSA) has compromised or got around the encryption techniques on which the security of Internet communications, electronic health records, e-commerce and banking are based.

The allegations — the latest in a series of disclosures about NSA activities — and Green's analysis of them should make one sit up and listen. "Not only does the worst possible hypothetical I discussed appear to be true, but it's true on a scale I couldn't even imagine," he wrote.

NSA mathematicians have been among the leading contributors to encryption research and the development of standards meant to protect the security of the Internet, often working closely with academic researchers and key bodies such as the respected US National Institute of Standards and Technology (NIST). But according to allegations made by *The New York Times*, *The Guardian* and public-interest-journalism website *ProPublica* in early September, based on documents provided by NSA whistle-blower Edward Snowden, the agency has also worked to weaken or create vulnerabilities in encryption standards.

Other allegations include collaborating with technology companies to provide entry points into their products, as well as forcing Internet companies to hand over encryption keys or hacking these from servers to defeat security.

On 9 September, NIST took the unprecedented step of opening a review of two of the suspect standards, and went so far as to warn users not to apply one of the standards until vulnerabilities had been double-checked by cryptographers. The Internet Engineering Task Force (IETF), an open international body that develops the core standards of the Internet, is now looking at how it can harden Internet protocols to

reinforce security and privacy against NSA-type attacks, and will take up the issue at its next meeting, in Vancouver, Canada, in November. This week, cryptographers at the University of Bristol, UK, published an open letter that called for a parliamentary enquiry into how security has been compromised.

Just as toxic subprime loans in the mid-2000s poisoned trust among financial institutions, leading to the financial crisis, the NSA's actions have poisoned people's trust in all the groups that make up the Internet ecosystem, from the giants of Google and Yahoo to telecoms companies, cloud-computing providers and the makers of chips and routers. US technology companies are likely to be the first to suffer, but the NSA's actions have corrupted the very fabric of the Internet.

Writing in *The Guardian*, cryptography researcher and security expert Bruce Schneier has called for scientists and engineers to take back the Internet, and for more whistle-blowers to come forward to detail how the NSA and authoritarian states are sabotaging electronic freedoms.

Certainly, mathematicians in the NSA, and external academics working with the agency, should examine their consciences. Mathematical associations and universities with links to the NSA should be more public and vocal about the revelations.

Like the IETF, Schneier wants scientists to re-engineer the Internet to make it more secure. Some technical improvements can be made — open-source code, which can be reviewed by anyone, is likely to be a major benefactor and facilitator — and the trust and security paradigms of developing Internet protocols have without doubt been irreversibly changed. But the Internet was not designed to be secure, and as the IETF points out on its blog, the scale of the NSA attacks was "not envisaged during the design of many Internet protocols".

As Schneier and the IETF acknowledge, technology is only part of the solution. Regulation of surveillance on the Internet and attacks on civil liberties are as much, or more, a question of policies. It has become abundantly clear over the past few months that there is but a fig leaf of oversight to protect against abuse of civil liberties by the NSA. The balance between security and civil liberties has gone off the charts in the wrong direction. ■

# Book smart

Novelist Thomas Pynchon shows that science and art can combine, with mutual benefit.

Some writers use metaphors in science. Some go further and make a metaphor of science itself — not the practical art of observation and empirical testing, but the often-tricky concepts at the heart of the pursuit. Such writing is difficult, and scientists and non-scientists alike can struggle with the result. But when done well, the language of research and the grammar of the natural world can sing a song as sweet as anything in literature. The supposed differences between the two cultures dissolve, leaving only those who get it and those who do not.

Many of those who do — both scientists and non-scientists — will be eagerly awaiting the latest book from Thomas Pynchon, *Bleeding Edge* (Penguin). It is reviewed on page 312 by Sean Carroll, a theoretical physicist at the California Institute of Technology in Pasadena who is himself a writer and a self-confessed Pynchon fan. Set against the terrorist attacks of 11 September 2001 in New York, *Bleeding Edge* is one of Pynchon's more straightforward books. As Carroll notes, it is "told linearly, from the point of view of an acknowledged main character, with something approximating an explicit goal".

That itself is a description of science, albeit a misleading one. Despite the appeal of a simple narrative, of cause and effect, and the dogged pursuit of truth by heroic individuals, most *Nature* readers will know — and no doubt lament — that science is not like that. Pynchon knows that too, and revels in our attempts to impose order on a chaotic, unruly reality.

Pynchon, Carroll notes, often uses imagery and symbolism from science and engineering. Stephen Hawking says that he was told that each equation printed in a popular-science book would halve its readership, so imagine the reaction of the editor on receiving the manuscript of Pynchon's 1973 classic novel *Gravity's Rainbow* (Viking), complete with a description of the first elements of the Poisson distribution. Organic chemistry, behaviour modification, double integrals and rocket dynamics all underpin both that story and the language that Pynchon chooses to tell it.

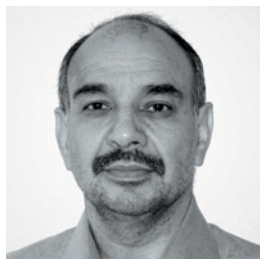
Some physicists consider Pynchon one of their own. The author studied for (but never finished) a degree in engineering physics at Cornell University in Ithaca, New York, and worked as a technical writer for the aerospace company Boeing. Biologists have credited his idea of a 'counterforce' — an organizing principle (also known as life) that counters the universal descent into entropy — as the spark that ignited their careers.

Those who get it see something special in Pynchon's work. There are few novelists who can claim to successfully unite the two cultures, but Pynchon does it by dispensing with metaphor and turning to science itself. ■

➔ **NATURE.COM**  
To comment online,  
click on Editorials at:  
[go.nature.com/xhnuq](http://go.nature.com/xhnuq)



NEIL PANDIT



## The Himalayas must be protected

*Climate change and human activities are pushing the fragile ecosystem ever closer to instability, warns Maharaj K. Pandit.*

As the Intergovernmental Panel on Climate Change (IPCC) gets ready to publish the first instalment of its latest report, many critics of the reports have been harking back to the previous effort — and the glaring mistake it made in stating that the Himalayan glaciers would disappear by 2035.

The situation in the Himalayas is not as dire as that, but it is certainly perilous. And although the IPCC overstated the timescale, the threat to the mountains from global warming and other pressures is genuine and must be addressed. Just this summer, the Indian Himalayan state of Uttarakhand witnessed what many described as the most horrific devastation in human memory. Cloudbursts, flash floods, landslides, human deaths and destruction were unleashed within a day. Human casualties were put in the thousands, and estimates of economic loss range from US\$500 million to around \$2 billion. Similarly, a monsoon flood hit Pakistan in 2010, killing more than 2,000 people, displacing millions and costing \$40 billion.

Policy-makers in India and elsewhere are reluctant to accept a glaring and dangerous truth: combined human activities have stressed the Himalayas close to their limit. We need positive action, not least a restriction on the number of tourists who visit during the monsoon season from June to September. Too many of them are unaware of the risks they run. Thousands of visitors perished in the Uttarakhand disaster. Policy-makers must actively engage with scientists and experts on the problems facing the Himalayas and their people to make sustainable development work.

Global warming can grab the headlines, but many of the other pressures on the fragile mountain region are more mundane. The human population is increasing fast in the Himalayas, and so is the speed of the landscape changes needed to support it. Cattle grazing and rampant deforestation — on current trends, one-third of the total Indian Himalayan forest cover could be gone by 2100 — will drive nearly one-quarter of endemic species to extinction and disrupt the natural flow of water. A changed Himalayan landscape caused by human activities and warming means transformed natural ecosystems through biological invasions and reduced native biodiversity.

People need energy, and in the Himalayas that means hydropower and dams. Nearly 400 new dams are proposed for the Indian and Tibetan Himalayas in the next decade or so. With them will come more change to ecosystems and more people displaced from submerged settlements. Such rampant dam-building in a region with high seismic activity and fragile geology shows that the policy-makers who approve these schemes either do not understand the scientific evidence or choose to ignore it.

The Himalayas are just 45 million years (Myr) old — mere striplings compared with the

Aravallis in India (around 4,000 Myr old) and the North American Appalachians (440–480 Myr). Young, folded and still rising, the Himalayas are more tectonically active than most mountains.

Rising temperatures add to the problems. Melting ice and snow form new glacial lakes, as well as increasing the volumes of existing ones. This could raise the threat of glacial-lake outburst floods. Some 8,800 glacial lakes in the Himalayas are spread across nations, and more than 200 of these have been classified as dangerous. Recent scientific evidence suggests that floods originating in the Himalayas are caused largely by landslides that temporarily block mountain rivers.

The Himalayas are warming faster than other mountain ranges, and the increased use of reinforced concrete in building construction, replacing the traditional wood and stone masonry there, is likely to create a heat-island effect and thus add to regional warming.

What is the way forward in the Himalayas? Clearly, the social and economic development of the Himalayan population cannot be undermined — literacy levels and school enrolment are up and infant mortality is down. Still, these are the same people who will suffer from the region's growing ecological degradation and environmental instability. Numerous regulations that should protect them, on mining and flood-plain development, for example, are poorly or rarely enforced. Indeed, many in India blame environmental regulations for the current economic downturn. An environmental tax on tourists to restrict numbers and raise funds would be equally unpopular, but the idea deserves proper discussion.

Most urgently, Himalayan countries need to build an international network that will monitor risks such as those from glacial lakes, and give early warning of hazards — similar to the tsunami warning systems installed around the Indian Ocean over the past decade. Scientists and engineers must make the case more forcefully that rampant building construction cannot be permitted on riverbanks or flood plains that are constantly swept by monsoon floods.

If the people of the Himalayas were more aware of the geological vulnerability and ecological fragility of their mountain home, they would surely force more compliance of laws and regulations to protect it. India and other affected countries should include in their school curricula basic knowledge of the geology and ecology of the Himalayas. If students are taught about their environment, they will feel more connected to the land and be more aware of its pulse. ■

**Maharaj K. Pandit** is a professor in the Department of Environmental Studies and director of the Centre for Interdisciplinary Studies of Mountain & Hill Environment, University of Delhi, India.  
e-mail: rajk.pandit@gmail.com

**HIMALAYAN  
COUNTRIES  
NEED TO BUILD AN  
INTERNATIONAL  
NETWORK  
TO MONITOR RISKS.**

➔ **NATURE.COM**  
Discuss this article  
online at:  
[go.nature.com/hlcrgrk](http://go.nature.com/hlcrgrk)

# RESEARCH HIGHLIGHTS

Selections from the  
scientific literature

## PHOTONICS

### Graphene sees the light

Three independent teams have boosted the performance of photodetectors that use graphene (atomically thick sheets of carbon) to convert light into electrical signals. This may enable fast optical communication within and between computer chips.

The latest devices boast data-transmission rates that rival those of conventional components and can sense a broader range of light wavelengths. Earlier graphene models produced weaker currents because most of the light failed to interact with their carbon layers. All three groups solved that problem by channelling light through silicon waveguides running along the graphene sheet.

The device built by Thomas Mueller at the Vienna University of Technology and his team produces 50 milliamps of current per watt of infrared light — only ten times less than germanium photodetectors, the current standard.

Dirk Englund of the Massachusetts Institute of Technology in Cambridge and his co-workers made a similar detector that handles an impressive 12 gigabits of data per second, and produces a clean electrical signal (pictured).

Meanwhile, a device from the group of Xiaomu Wang, then at the Chinese University of Hong Kong, picks up mid-infrared light at room



## ECOLOGY

### Another killer fungus hits amphibians

The Netherlands' population of wild fire salamanders (*Salamandra atra*; pictured) declined by 96% in the past three years, but no known infectious agent was found on their bodies. Now An Martel at Ghent University in Merelbeke, Belgium, and her team identify the problem as a new species of chytrid fungus, *Batrachochytrium salamandrivorans*. Healthy salamanders that were experimentally infected with the fungus developed skin lesions and died.

Unlike the only other chytrid fungus known to

cause deadly infections (*B. dendrobatidis*, which has ravaged global frog and toad populations), this new species does not affect midwife toads (*Alytes obstetricans*). It also grows at much lower temperatures, suggesting that the two chytrid species occupy different niches. The researchers developed a DNA-testing method to rapidly screen salamanders for the fungus, with the aim of tracking this latest threat to biodiversity.

**Proc. Natl Acad. Sci. USA** <http://doi.org/nrz> (2013)

temperature, suggesting that graphene detectors could bypass the restrictive cooling requirements of other photodetector materials.

**Nature Photon.** <http://doi.org/ns8>; <http://doi.org/ns9>; <http://doi.org/ntb> (2013)  
For a longer story on this research, see: [go.nature.com/zoioiw](http://go.nature.com/zoioiw)

## NEUROSCIENCE

### Different strains of Alzheimer's

Protein fibres that build up in the brains of people with Alzheimer's disease take on different structures in

different patients.

Aggregates of the amyloid- $\beta$  protein are a hallmark of Alzheimer's disease, but the molecular forms that they take in the brain have not been explored. Robert Tycko at the US National Institutes of Health in Bethesda, Maryland, and his team extracted amyloid- $\beta$  from the brains of two patients who had died, and who had displayed different symptoms.

They used this amyloid- $\beta$  to seed the growth of synthetic fibrils and analysed each sample with nuclear magnetic resonance and electron microscopy. Unexpectedly, all

the fibrils from each brain had the same molecular structure, but the structures differed between brains.

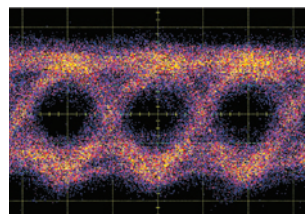
The researchers say that imaging agents that distinguish between fibril structures might allow more-precise diagnoses of Alzheimer's disease.

**Cell** 154, 1257–1268 (2013)

## STRUCTURAL BIOLOGY

### Rigid receptor denies HIV entry

The molecular structure of a protein commandeered by HIV to enter human cells reveals sites that could lead to





better anti-HIV drugs.

The protein CCR5 resides on the surface of certain immune cells. A team led by Beili Wu at the Chinese Academy of Sciences' Shanghai Institute of Materia Medica successfully crystallized CCR5 bound to the anti-HIV drug maraviroc. The structure suggests that the drug works by making CCR5 more rigid, and so prevents movements that allow HIV to enter cells. It also reveals why such HIV drugs fail to work with certain CCR5 mutants.

Differences between CCR5 and CXCR4, another protein used by HIV to enter cells, may explain why some HIV strains favour one protein gateway over the other.

**Science** <http://doi.org/ntj> (2013)

## ZOOLOGY

## Fertility smells like preen spirit

A bird's scent may indicate how many offspring it will produce.

Danielle Whittaker of Michigan State University in East Lansing and her colleagues analysed compounds that evaporate from the oily secretions that birds spread over their feathers when preening. The team collected oil from 12 female and 22 male dark-eyed juncos (*Junco hyemalis*; pictured) and found that the oil's chemical profile differed between the sexes.

Birds that released more chemicals characteristic of their sex produced more offspring. And males with more 'female' odours fledged more hatchlings fathered by other birds from their nests. Overall, the smell of a bird was a better predictor of reproductive success than either size or plumage.

**Anim. Behav.** <http://doi.org/nr3> (2013)

## NATURAL PRODUCTS

## African tree gets to the root of pain

An African plant used in traditional medicine for pain relief contains the same active ingredient as an artificial pain killer.

Together with scientists in Cameroon, France and Switzerland, Michel De Waard at Joseph Fourier University in Grenoble, France, collected extracts from the pincushion tree (*Nauclea latifolia*), separated compounds into groups on the basis of their mass and chemical properties, then tested each group in mice for its ability to relieve pain.

The team found an oily yellow compound in the most potent group and determined its chemical structure to be that of tramadol, which has been sold as a synthetic analgesic since the 1970s. The compound was detected only in the plant's roots — the same parts used in traditional remedies to treat pain. The researchers believe that this is the first time a widely prescribed synthetic drug has been found in a plant at clinically relevant concentrations.

**Angew. Chem.** <http://doi.org/f2dv27> (2013)

## GENETICS

## Ancient bear bone yields a sequence

The mitochondrial genome of a Pleistocene cave bear (*Ursus deningeri*) has been reconstructed using extremely short DNA molecules from a bone that is more than 300,000 years old. Apart from rare specimens preserved in permafrost, the fossil is some 200,000 years older than any other material used to generate a complete DNA sequence.

By reworking methods to purify the tiny amounts of damaged DNA that are typical of old samples,

## COMMUNITY CHOICE

The most viewed papers in science

## NUTRITION

## Fruits shrink diabetes risk

**HIGHLY READ**  
on [www.bmj.com](http://www.bmj.com)  
in August

Three large, long-running health surveys suggest that some fruits — but not fruit juice — reduce the risk of diabetes.

Researchers led by Qi Sun at the Harvard School of Public Health in Boston, Massachusetts, looked at people's reports of how often they ate certain fruits and whether they developed diabetes. Results came from a total of 187,382 US health professionals whose health and habits have been tracked for more than two decades, during which time more than 12,000 of the study participants developed diabetes.

Eating whole fruits such as apples, pears, grapes, raisins and particularly blueberries was associated with a lower risk of diabetes. By contrast, drinking fruit juice was associated with a slight increase in risk, perhaps because juices contain fewer fruit-specific chemicals and introduce more sugar into the bloodstream more quickly.

**Br. Med. J.** 347, f5001 (2013)

Jesse Dabney at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany, and his colleagues collected and sequenced DNA strands as short as 30 basepairs. More than 90% of the sequences used to produce the genome were less than 50 basepairs long — too short to be used efficiently with conventional methods.

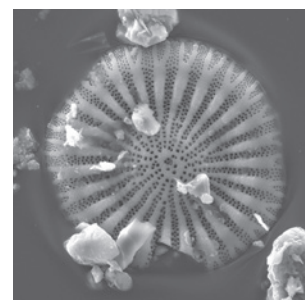
The technique could allow researchers to recover DNA from ancient humans and other specimens that are currently considered too degraded for most analyses. **Proc. Natl Acad. Sci. USA** <http://doi.org/nr4> (2013)

## VOLCANOLOGY

## Eruption sent microbes flying

Fragments of once-living creatures inside volcanic rocks can help to trace an eruption to its source.

New Zealand's North Island is blanketed in debris from a super-eruption of the Taupo volcano at the island's centre, which happened some 25,000 years ago. A team led by Alexa Van Eaton at the Victoria University



ALEXA R. VAN EATON/DAVID FLYNN

of Wellington searched for microfossils in rocks as far as 850 kilometres from the volcano. They found abundant skeletons of algae known as diatoms (pictured), including a type that lives only in lakes on the North Island. This confirms the findings of earlier work that the eruption blasted through a lake on that island.

Such fossils could help volcanologists to work out the locations and environmental settings of past eruptions, the authors say. They speculate that volcanoes might even disperse living cells across long distances.

**Geology** <http://doi.org/nr8> (2013)

**NATURE.COM**

For the latest research published by Nature visit:

[www.nature.com/latestresearch](http://www.nature.com/latestresearch)





# SEVEN DAYS

The news in brief

## POLICY

### Encryption crisis

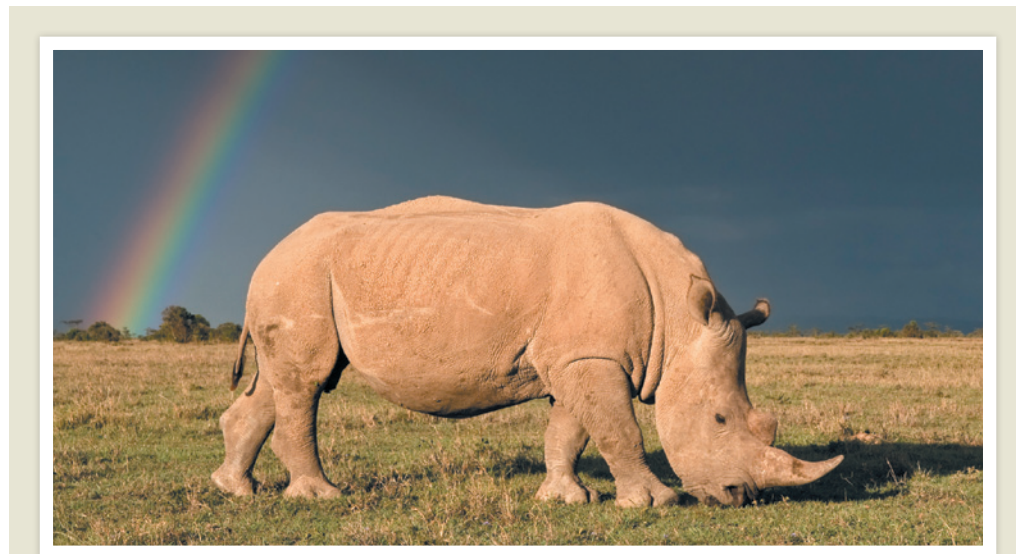
The US National Institute of Standards and Technology (NIST) was last week struggling to maintain trust among information scientists, after media reports that the US National Security Agency had compromised NIST-approved algorithms used to encrypt data. “NIST would not deliberately weaken a cryptographic standard,” the agency, headquartered in Gaithersburg, Maryland, said in a statement on 10 September. A day earlier, it had said that one of its approved algorithms should not be used because of a problem with how it generated random numbers. See page 282 for more.

### No science minister

Australia’s incoming government will do without a science minister — or so it seems from Prime Minister-elect Tony Abbott’s cabinet, announced on 16 September. Abbott avoided explicitly naming science in any minister’s portfolio (the first such omission since 1931). Much of the science portfolio would come under the industry ministry, he said.

### Nuclear restart

North Korea may be restarting its nuclear reactor, a 5-megawatt plant in Yongbyon that could produce plutonium for nuclear weapons. Satellite imagery suggests that a building near the reactor is venting steam, reported analysts on the website 38 North ([go.nature.com/pgcyqf](http://go.nature.com/pgcyqf)), a programme of the US–Korea institute at the Paul H. Nitze School of Advanced International Studies in Washington DC. North Korea had said in April that it would restart the reactor, which



MARK CARWARDINE/NATUREPL.COM

## More protection for endangered rhinos

The US Fish and Wildlife Service has added the southern white rhinoceros (*Ceratotherium simum simum*, pictured) to the list of animals protected under the US Endangered Species Act, to fight an increase in poaching abroad. The sub-species was already protected under international rules on trade in wildlife, but

sales of hunting trophies were allowed. This left a loophole under which horns from more endangered — and better protected — animals such as the northern white rhino (*Ceratotherium simum cottoni*) were passed off as trophies from southern white rhinos. The new rule increases restrictions on horn sales in the United States.

was shut down and partly demolished in 2008 as part of now-stalled disarmament talks.

### Chemical weapons

Syria’s agreement to join the Chemical Weapons Convention last week means that only six states — Egypt, Israel, Angola, North Korea, South Sudan and Myanmar — have not joined the treaty. Russia and the United States were waiting this week to see whether Syria would disclose its chemical-weapons arsenal and related facilities, a necessary step before the weapons can be destroyed. There is “clear and convincing evidence” that rockets containing the nerve gas Sarin were used in Damascus on 21 August, United Nations weapons inspectors said on 16 September.

### Biofuels overhaul

The European Parliament has voted to limit crop-based biofuels, after criticisms that fuels such as biodiesel from palm oil can produce greater greenhouse-gas emissions than fossil fuels, and can contribute to deforestation and rising food prices (see *Nature* **499**, 13–14; 2013). The new policy retains a target that 10% of transport fuels should come from renewable sources by 2020 — but sets a 6% cap on crop-based biofuels. However, another vote will be needed because some politicians objected. See [go.nature.com/yqzrug](http://go.nature.com/yqzrug) for more.

### Laureate scuttled

A proposal to create a ‘US science laureate’ position has hit a roadblock. The US House of Representatives cancelled a

planned 10 September vote on a bill to establish the unpaid, honorary position after a conservative group objected to the plan as wasteful and unnecessary. The bipartisan group of lawmakers behind the laureate plan says that it will hold a hearing on the measure next month.

## RESEARCH

### Voyager breakaway

The Voyager 1 spacecraft has entered interstellar space 36 years after its launch, physicists reported on 12 September (D. A. Gurnett *et al.* *Science* <http://doi.org/nsv>; 2013). According to plasma-density measurements, the spacecraft left the heliosphere — the Sun’s cocoon of charged particles — in late August 2012. The findings settle an argument

that has raged among Voyager team members for more than a year (see *Nature* **489**, 20–21, 2012). The probe remains short of the Oort cloud, a reservoir of comets that is still part of the Solar System. See [go.nature.com/cdkxck](http://go.nature.com/cdkxck) for more.

## Meningitis triumph

A cheap vaccine against meningitis A has proved a huge success, boosting hopes of cleaning up Africa's 'meningitis belt' — where the disease periodically kills thousands in intense epidemics. Mass vaccination of 1.8 million people in Chad cut cases of meningitis by 94% relative to unvaccinated parts of the country, researchers reported on 12 September (D. M. Daugla *et al.* *Lancet* <http://doi.org/nr6>; 2013). The GAVI Alliance in Geneva, Switzerland, will have rolled out the vaccine in 12 countries by the end of 2013, and has committed US\$267 million to reach another 14 countries by 2016.

## Kenya strikes water

Massive reserves of groundwater have been discovered in parched northern Kenya. Exploration firm Radar Technologies International (RTI) used satellite and seismic data to find aquifers that contain at least 250 billion cubic metres of water, it announced on 11 September. Confirmed by drilling



(pictured), the renewable part of the water reserve (replenished by rainfall each year) amounts to 3.4 billion cubic metres, boosting Kenya's renewable water resources by 17%, RTI says.

## HeLa data approval

An advisory committee within the US National Institutes of Health gave researchers permission to use genomic data from the HeLa cell line for the first time on 16 September. The announcement follows several months of negotiations over data-use rights with the family of Henrietta Lacks, whose cervical tumour gave rise to the HeLa cell line in 1951. See [go.nature.com/xbscjv](http://go.nature.com/xbscjv) for more.

## UK trial registry

Medical researchers in the United Kingdom will need to register their clinical trials in a publicly accessible database to gain ethical approval, under

a policy that comes into effect from 30 September. The Health Research Authority in London announced the policy on 11 September. Compulsory registration has been a key demand of campaigners for greater openness in medical research; some trials must already be registered in some jurisdictions, but compliance has been patchy.

## Brain roadmap

The US National Institutes of Health in Bethesda, Maryland, has revealed nine broad research goals for its ambitious, multi-year brain-mapping initiative, which was announced in April by US President Barack Obama. Priorities include improving knowledge of the brain's cell types, how they are connected and how they communicate, says the agency's 16 September report. The plan will be further refined up to June 2014. See [go.nature.com/iyrueq](http://go.nature.com/iyrueq) for more.

### FACILITIES

## Flash flooding

Unusually heavy rains have flooded more than 6,100 square kilometres of Colorado. Between 11 and 17 September, rising waters destroyed roughly 1,500 homes in the state's most populated region, killed at least 8 people and displaced 11,750 more, according to the

## COMING UP

### 21–24 SEPTEMBER

The European Molecular Biology Organization meets in Amsterdam. Discussions will include the future of cancer genomics. <http://the-embo-meeting.org>

Colorado Office of Emergency Management. The floods also closed major scientific facilities in Boulder, including the University of Colorado and the National Center for Atmospheric Research, for several days.

### BUSINESS

## Fracking laws

California passed a bill on 11 September that puts strict requirements on companies that produce oil and gas through hydraulic fracturing ('fracking'). State governor Jerry Brown is expected to sign the law, which requires firms to seek permits before fracking, and to disclose to the state all of the chemicals they use. Southern California's San Joaquin and Los Angeles basins hold 64% of the technically recoverable US shale-oil reserve, according to the US Energy Information Administration.

## Sclerosis success

The monoclonal antibody alemtuzumab was approved to treat multiple sclerosis (MS) by the European Commission on 17 September. The immune-quieting antibody was first developed in the 1980s, and found use in treating lymphomas and preventing rejection of organ transplants. The drug was developed for MS treatment by Genzyme, based in Cambridge, Massachusetts. See [go.nature.com/phawy5](http://go.nature.com/phawy5) for more on alemtuzumab's clinical-trial success.

➔ [NATURE.COM](http://NATURE.COM)

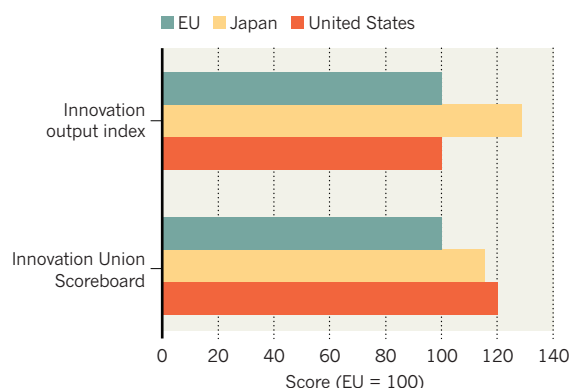
For daily news updates see: [www.nature.com/news](http://www.nature.com/news)

## TREND WATCH

Each year, the European Union's 'innovation scoreboard' puts the region behind the United States on an index based on a raft of measures, including scientific publications, research spending and patents. But on 13 September, the European Commission announced a new index focusing on innovation output — broadly speaking, how well nations squeeze economic value out of their ideas. This scores the regions as equal. The two indicators will be published alongside each other in future.

## TRANSATLANTIC INNOVATION RACE

A new measure of innovation output puts the European Union (EU) on level pegging with the United States, unlike the old Scoreboard.





# NEWS IN FOCUS

**AGRICULTURE** Greener grass bred to cut planet-warming gases from soil **p.291**

**ASTRONOMY** Ripples spread from labour strike at Chile's ALMA telescope **p.292**

**INFECTIOUS DISEASE** One year on, we know far too little about MERS coronavirus **p.294**

**CLIMATE CHANGE** As the IPCC turns 25, *Nature* explores its evolution and impact **p.297**

THOMAS PETER/REUTERS/CORBIS



German Chancellor Angela Merkel has her eyes on a third term in office.

POLICY

## Germany hits science high

*But election prompts fears that budgetary pressures may sap strong investment.*

BY QUIRIN SCHIERMEIER

As physicist by training, German Chancellor Angela Merkel has a proven affinity for science. Thanks to generous government funding, German research has thrived since she first came to power eight years ago. But as she prepares for a probable third term in office after elections on 22 September — the governing coalition between her Christian Democratic Union (CDU) party and the Free Democrats is well ahead in polls — scientists and science organizations are concerned that the years of plenty may soon be over.

On the face of it, there does not seem too much to worry about. Since 2005, the federal government's overall science spending has increased by a whopping 60% — from €9 billion (US\$12 billion) up to around €14.4 billion in

2013 (see 'The rise of German science'). By comparison, in 1995–2005, the German science budget rose by only 7.5%. Industrial research has also flourished: Germany is now close to spending 3% of its gross domestic product (GDP) on science and technology, a key target of the European Union (EU) 2020 strategy for growth that only Finland, Sweden and Denmark have met so far.

These investments have paid off. This month, the World Economic Forum, based in Switzerland, moved Germany up two notches on its ranking of the world's most competitive economies, noting that heavy investment in research and development has added to Germany's strengths. It now stands at number four, behind Switzerland, Singapore and Finland.

"It's a great achievement of Merkel's chancellorship to have firmly anchored science as an

important pillar of our country's wealth and innovative power," says Peter Gruss, outgoing president of the Max Planck Society, which runs 82 research institutes in Germany. But he fears that the good times may end if other budget pressures, such as debt reduction, welfare and transport, are prioritized after the election.

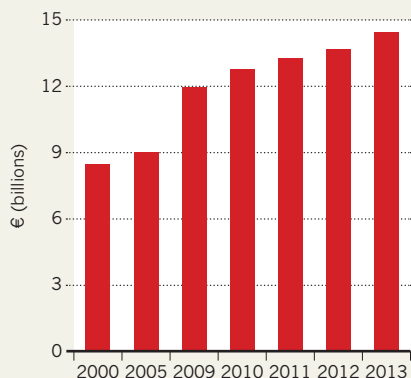
The German political landscape is complex. The ruling coalition's main competitors are Peer Steinbrück's left-leaning Social Democrats. But they, too, would need a coalition partner to govern. The most likely candidates are the Greens, but such junior parties often fail to win the 5% of votes needed to secure parliamentary representation, so a grand coalition of the CDU and Social Democrats cannot be ruled out.

The scientific landscape is more clear-cut. In their election programmes, all parties represented in the current parliament — including ►



## THE RISE OF GERMAN SCIENCE

Since 2000, German science spending has increased by about 70% and is now approaching 3% of GDP.



► the Greens and the Left party — have taken a decidedly pro-science stance. They all, for example, see energy research as a priority to support Germany's planned transition, known as the *Energiewende*, to a non-nuclear, low-carbon system dominated by renewable energy sources (see *Nature* **496**, 156–158; 2013). Likewise, there is cross-party agreement that health research must be strengthened and remodelled, a policy driven by Germany's rapidly ageing population (see 'Health considerations').

All parties claim that climate change is another priority, and none will axe any commitments already made on major science issues, such as Germany's decision to phase out nuclear power by 2022.

For her part, Merkel has vowed not to "fall behind" the 3% GDP spend on research and development. But her government, burdened by debt and aid payments to Greece and other struggling eurozone countries, has set itself tough fiscal restrictions. A 2009 constitutional amendment called the *Schuldenbremse* ('debt brake') means that from 2016, the federal government — and from 2020, state governments — will no longer be allowed to run deficits.

For the moment, the science boom continues. Over the past eight years, the large German science organizations — including the Max Planck Society, the Helmholtz Association of German Research Centres and the German Research Foundation (DFG), the main funder of university research — have benefited from the Pact for Research and Innovation, which

they jointly negotiated with the two previous governments to give institutions financial-planning security. Each organization received 3% annual budget rises between 2005 and 2008, and 5% annual increases since. And the science ministry is contributing €3.4 billion to the Excellence Initiative for universities, a €4.6-billion government scheme that has been used to set up graduate schools and research collaborations (see *Nature* **487**, 519–521; 2012).

For the DFG, the boom has had mixed effects. Despite its bigger coffers, the success rate of applications for individual DFG grants dropped from 47% in 2009 to 32% last year. This is in part because more scientists are applying for grants, and because grant size has increased. It means that some excellent proposals will not get funded, says DFG president Peter Strohschneider. He worries that the falling success rate could demotivate researchers.

But by and large, the mood is upbeat. German science has shed its narrow outlook over the past decade. At Max Planck institutes, almost 90% of postdocs, half of all postgraduate students and more than 40% of scientific directors recruited in the past decade came from abroad. And the society has created cutting-edge institutes in emerging fields. In Mühlheim, scientists are exploring ways to improve chemical energy conversion and energy storage. And ocean scientists are looking forward to the completion of a €450-million icebreaker that is to replace the ageing flagship of Germany's research fleet.

Not everything is rosy, however. The ruling

unifies local health research at different sites in the city and also receives local state funds.

There are two options to smooth out such anomalies: turn the national centres into a German version of the US National Institutes of Health, or build up networks of local health institutes.

The government will therefore need to decide whether to focus on "localized cathedrals, like the Berlin Institute of Health, or creating more national institutes", says Pierluigi Nicotera, head of the DZNE.

## GERMAN RESEARCH CENTRES

## Health considerations

One of the first issues facing the next government will be streamlining health-research facilities. Currently a patchwork of locally and nationally funded centres, the network is crying out for more coherence.

Since 2009, six national health-research centres have been created, including the German Center for Neurodegenerative Diseases (DZNE) in Bonn. But the federal government is also, for example, injecting some €300 million (US\$400 million) into the Berlin Institute of Health, a hybrid body that

coalition parties and the Social Democrats are in favour of continuing the Pact for Research and Innovation beyond its planned 2015 expiry, but this is by no means guaranteed. And the Excellence Initiative finishes in 2017, raising fears that thousands of scientists employed on fixed-term contracts might be rendered jobless.

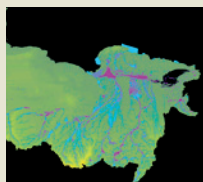
Junior researchers are especially at risk. Despite Germany's strong research presence, the country "needs to improve the career development of young scientists — there is not enough tenure-track here", says Oliver Brüstle, a stem-cell researcher at the University of Bonn. In Germany's federal system, universities fall under the governance of the 16 states. As student numbers continue to rise, many institutions, especially in the poorer states in the north and east, are suffering from relative underfunding. A proposed constitutional change that would have allowed the central government to fund universities directly, and boost salaries of lecturers and researchers, last year failed to win support among the states. All parties say that they will tackle the issue again after the elections.

"For all the money that has been injected into the system, universities cannot offer more-permanent positions if they don't know if the cash will keep flowing," says Horst Hippler, president of the German Rectors' Conference. "The next government must urgently remedy that matter — or else the boom will have been a bubble." ■

Additional reporting by Alison Abbott.

## MORE ONLINE

## TOP STORY

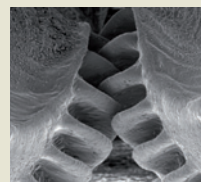


Ancient rivers cut migration routes through Sahara  
[go.nature.com/rknpqf](http://go.nature.com/rknpqf)

## MORE NEWS

- Did a hyper-black hole spawn the Universe? [go.nature.com/wtjfiip](http://go.nature.com/wtjfiip)
- European Research Council funds arXiv [go.nature.com/8xpkx7](http://go.nature.com/8xpkx7)
- Ancient mantle plume left scar on North American continent [go.nature.com/ekrwe8](http://go.nature.com/ekrwe8)

## VIDEO OF THE WEEK



Leg cogs help insects to synchronize limbs for jumping  
[go.nature.com/6gig5d](http://go.nature.com/6gig5d)



*Brachiaria* grasses, used as cattle pasture, have been bred to reduce nitrogen emissions caused by microbes.

## AGRICULTURE

# Grass gets greener

*Plant secretion curbs greenhouse-gas emissions from soil.*

BY NICOLA JONES

Industrial power plants take most of the flack for climate change. But nature's plants are not blameless: agricultural land is responsible for about 14% of the world's greenhouse-gas emissions, slightly more than the global contribution from planes, trains and automobiles. So researchers have tried to 'green' up agriculture by breeding plants that cut the emissions associated with wasted fertilizer.

On 13 September, researchers announced that they have bred a tropical pasture grass that can significantly suppress greenhouse-gas emissions. The team, from the International Center for Tropical Agriculture (CIAT) in Cali, Colombia, is working with Dow AgroSciences, based in Indianapolis, Indiana, to get seeds onto the market in the next 3–5 years.

Agriculture's climate problem is a nitrogen-fertilizer problem. Fertilizer contains ammonium ( $\text{NH}_4^+$ ); when it is first laid down, this positively charged ion stays put in the soil, sticking to negatively charged clay particles. But then nitrifying soil bacteria go to work, wreaking environmental havoc. They convert ammonium to nitrate ( $\text{NO}_3^-$ ), which washes into ponds and causes ecologically harmful algal blooms. Nitrate can also be converted to nitrous oxide ( $\text{N}_2\text{O}$ ) gas, which warms the planet 300 times more powerfully than carbon

dioxide. Less than one-third of the nitrogen applied as fertilizer typically makes it into crops.

Rising food demand is driving increased fertilizer use, and the problem is expected to get worse: nitrous oxide emissions will be 50% higher in 2020 than in 1990, according to the US Environmental Protection Agency. "The nitrogen dilemma is a huge issue," says Henry Janzen, a soil biochemist at Agriculture and Agri-Food Canada in Lethbridge, and co-author of the chapter about agriculture in a 2007 report on climate mitigation by the Intergovernmental Panel on Climate Change. "It's a massive undertaking to reduce these emissions while making more food."

## SLOW DOWN

The solution, says Michael Peters, an agronomist at CIAT and leader of the team that has developed the low-emissions grass, is to encourage ammonium to persist in the soil for longer, by suppressing microbial activity. Farmers can buy synthetic nitrification inhibitors such as dicyandiamide, but these are not ideal. The chemicals can be washed away, and it would be impossible to target them to where they are needed most — where grazing animals have left urine and dung that act as extra fertilizer.

In the 1980s, CIAT researchers noticed that some grasses grow well even without fertilizer — particularly *Brachiaria humidicola*, which

is adapted to low-nitrogen South American savannahs. After years of hunting, they identified a nitrification inhibitor secreted by the grass's roots. Called brachialactone, it reduces nitrous oxide emissions by blocking enzymatic pathways in nitrifying bacteria. That leaves more nitrogen available to help the plant to build tissues.

The team found similar activity in a few other plants, including sorghum, but *Brachiaria* grasses were best. The researchers have spent more than 8 years breeding the plants to maximize this ability. Peters says that they have doubled the release of nitrification inhibitors, and are now checking that this has not decreased the overall productivity of the grass. As a side-benefit, the team reported this week at the 22nd International Grasslands Congress in Sydney, Australia, brachialactone seems to persist in the soil: wheat grown in pastures that previously hosted *Brachiaria* produces nearly four times as much grain with low levels of fertilizer as wheat grown in previously cropped land.

"The idea has legs," says Peter Grace, a soil scientist at the Queensland University of Technology in Brisbane, Australia, and coordinator of Australia's National Agricultural Nitrous Oxide Research Programme.

## CROP CHALLENGE

Attacking the problem in grasslands is important: 85% of Earth's agricultural land is used for livestock forage. But crops, which are more heavily fertilized than grasslands, are much worse climate offenders. To tackle that problem, CIAT geneticists are trying to isolate the brachialactone genes, to introduce them into crops such as wheat or rice. Plants could then produce their own inhibitors when they sense high concentrations of ammonium in the soil.

Help might also come from genes that encourage faster uptake of nitrogen. Arcadia Biosciences, based in Davis, California, has taken a gene for an enzyme called alanine aminotransferase from barley and incorporated it into other crops to encourage them to absorb nitrogen before microbes do. On 10 September, Arcadia announced that in trials by CIAT, African rice with this gene produced the same yield as controls, using only half the fertilizer. Arcadia expects commercial partners to bring seeds to market by 2017, and in December 2012, the United Nations Clean Development Mechanism approved a plan for farmers to earn sellable 'emissions reduction' credits by using the company's technology.

Grace points out that there are also effective low-tech approaches to reducing emissions. Farmers can refine how and when they apply fertilizers; for example, they can avoid lots of nitrogen being washed away when snow melts. Tilling less and periodically planting crop fields with nitrogen-fixing legumes can also help to keep nitrogen in the soil. "The easiest way is to simply restrict the amount of fertilizer applied," says Grace. ■





Striking workers at the ALMA telescope array occupied the operations building for more than two weeks.

# ASTRONOMY

## ALMA strike stirs up Chilean labour unions

*Other observatories watch for sharpened worker demands.*

BY ALEXANDRA WITZE

Messages of support raced across Chile's high desert plateaus. A labour strike had halted work at the world's premier radio telescope, the Atacama Large Millimeter/submillimeter Array (ALMA), and word was spreading to other elite astronomical observatories in the country.

Suddenly, the Chilean workers who staff these observatories — from the engineers who keep telescope software going to the mechanics who service broken equipment — were sharing the world stage. As the strike began at ALMA, workers more than 300 kilometres to the southwest at the Paranal observatory put down their radios instead of handling after-hours calls — not in a strike, but as a gesture of solidarity.

The 17-day dispute at ALMA ended on 7 September, when the labour union reached an agreement with the management that included shorter work shifts and pay rises for some of the 195 union members. But the walkout, the first unionized strike at an international observatory in Chile, underscores the complicated relationship between the predominantly US- and European-funded observatories and their Chilean workers. At Paranal, which is run by the European Southern Observatory (ESO), workers say that the

events at ALMA have emboldened them to speak out and ask for more concessions from management. "What the ALMA union did is open the road for every observatory in Chile to talk about what is going on," says Nicolas Slusarenko, president of the Paranal union and a software engineer at the observatory's Very Large Telescope.

The roots of many of the workers' complaints date back to the 1960s, when US and European astronomers first began to build telescopes in the high Atacama Desert, which is coveted for its pristine night-viewing conditions. Chile, ruled through the 1970s and 1980s by dictator Augusto Pinochet, was viewed as a politically stable place. Construction and labour costs were cheap in comparison to those at other prime sites, such as Hawaii, and Chile was eager to lure investment. Over the years, Chilean governments have granted privileges to certain observatory workers, such as income-tax exemptions for European astronomers working in Chile — but such privileges do not apply to Chileans.

In practice, this means that the observatories are run like small states within Chile — a fact that rankles some union leaders. Slusarenko says that whereas unsafe Chilean businesses could be subject to inspection by a labour ministry agency, no such avenue exists at ESO sites. "Where do I go to protect employees?" he asks.

But the unions have been gaining strength. In the past couple of years, the Paranal union has added members from three other ESO locations to form a 'sindicato' with 105 members overall. And in April 2012, a federation representing about 500 workers was created by joining the ESO and ALMA unions with two others representing workers at US-run observatories.

Last November, a walkout was narrowly avoided at Paranal. The union voted to go on strike over issues such as maternity leave, but that was averted by an eleventh-hour agreement with ESO. That contract expires in November 2014, says Slusarenko, and he expects renewal negotiations to be stressful. "It will be interesting to see if they take into account what happened at ALMA," he says. ESO management says that it has worked hard to develop good relations with its Chilean staff and does not expect any trouble. "We have no reason to expect any change to ESO operations as a result of the recent ALMA strike," says ESO director general Tim de Zeeuw.

Chris Smith, who oversees operations in Chile for the Association of Universities for Research in Astronomy (AURA), a US-based group that operates three sites, also downplays the effect of the ALMA strike. He says that there is a family-like atmosphere at the AURA observatories, and management meets with union representatives every month or two to deal with any labour issues before things get too tense. The ALMA strike is an anomaly, he says. "I don't want anyone to walk away thinking that this sort of conflict is the norm — it's not."

The ALMA strike began after the old labour contract expired and the ALMA union asked for changes in the new contract. The union wanted a 15% pay rise, plus other concessions that were standard at other observatories, such as a working week of 40 hours instead of 45 for administrative staff, and workdays of 11 hours instead of 12 for shift workers. Their employer, the US-based Associated Universities Incorporated (AUI), balked, and the strike began on 22 August.

Workers occupied the main operations building and paraded underneath ALMA's dishes with banners demanding justice. They produced a slick YouTube video (see [go.nature.com/jl7z3r](http://go.nature.com/jl7z3r)) with scenes from the strike. Science observations stopped completely, although some work continued — the project accepted delivery of its next-to-last antenna on 28 August, says Al Wootten, an ALMA project scientist. ALMA is a collaboration between, primarily, ESO, the US National Radio Astronomy Observatory (managed by AUI), and the National Astronomical Observatory of Japan.

After more than two weeks of negotiations, which included the personal intervention of the head of the Chilean labour department, the union and AUI signed a new two-year contract. The shorter working hours start in January, and employees who work at sites

MARCELO H. BARTSCH/INDICATO ALMA



above 5,000 metres will receive a bonus, but the only pay rise granted was 4% to the lowest-paid union members. Staff were paid for their strike days and received a bonus for ending the strike. The AUI would not say how much the extra concessions would cost; ALMA director Pierre Cox will decide how to pay for them

within the project's fixed budget.

The final telescope antenna in the 66-dish array is scheduled for delivery this month, and science observations by principal investigators are expected to resume in early October, says Ethan Schreier, the AUI's president. "We're looking forward to just getting back," he says. ■

## RESEARCH

# Israel in deadlock on Horizon 2020

*Conciliation sought in talks on EU research guidelines.*

BY ALISON ABBOTT

When cell biologist Benny Geiger learned that Israel may decide not to participate in the next multibillion-euro research programme of the European Union (EU), he was dumbfounded. "It seemed unreal — my first thought was that there must have been some sort of mistake," he says.

There was not. Israel had objected to EU guidelines published in July, which stated that any new agreement it made with the EU would apply only to territories within its pre-1967 borders. This would make universities and other institutions ineligible for funding of research set to take place in the occupied territories. On 14 August, in a formal response, Israeli officials charged that the policy was an attempt by the EU to push Israel to waive its claims of sovereignty over the territories. As a result, officials announced that the country would not sign up to Horizon 2020 — the

€80-billion (US\$107-billion) financial instrument designed to secure European competitiveness — under the proposed conditions.

For the Israeli scientific community, the stakes are very high. Not being part of Horizon 2020 would be "a huge blow on many levels", says Geiger, of the Weizmann Institute of Science in Rehovot, and chairman of the academic board of the Israel Science Foundation, the country's main research-funding organization. It is not just about the money, he says: it is also about the inability to integrate into the EU scientific community, and the potential loss of collaborative ties. Both would be bad news given Israel's geographical isolation in a region that otherwise has little research activity.

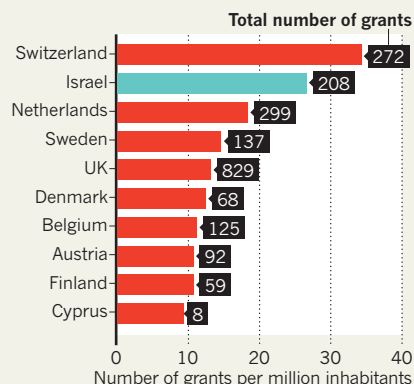
The roots of the conflict lie in the EU's policy of not recognizing Israel's sovereignty over territories it has occupied since 1967, including East Jerusalem, the Golan Heights, the Gaza Strip and the West Bank. In December last year, EU foreign ministers ruled that "all agreements between the state of Israel and the European Union must unequivocally and explicitly indicate their inapplicability to the territories occupied by Israel in 1967".

Israel has been involved in EU research programmes as an associate country since 1996, and is one of just a handful of non-EU member states — including Switzerland and Norway — to have this status. It has benefited enormously. The country paid just €534 million to join the EU's Seventh Framework Programme (FP7), the forerunner of Horizon 2020, but its net gain once FP7 finishes at the end of this year will be about €634 million, with almost 1,600 Israeli scientists participating in grant agreements under the programme.

Almost half of this money supports Israeli participation in various FP7 cooperation projects. The rest went to winners of European Research Council (ERC) grants. Israeli scientists have been unusually successful in ▶

## PUNCHING ABOVE ITS WEIGHT

Per capita, Israel has received the second-highest number of European Research Council (ERC) grants among European Union and associate members.



\*Data cover ERC starting grants for 2007–13 and ERC advanced grants for 2008–12.

SOURCE: EUROPEAN RESEARCH COUNCIL

► these, winning prestigious and highly competitive individual-investigator grants for basic-research projects. More than 30 were awarded ERC funding in its sixth Starting Grant competition in July, for example. The country is second only to Switzerland in its number of ERC grantees per capita (see 'Punching above its weight').

"We are watching with a great sadness — even horror — at the mixing of politics into science," says Geiger, himself a holder of an ERC grant. "There is a lot to lose."

In a bid to break the deadlock, another round of negotiations took place in Brussels last week. No concrete progress was made, but Michael Jennings, a European Commission spokesman, described the day-long talks as "constructive and open", with both sides hoping to find a resolution before Horizon 2020 launches on 1 January.

He added that the restriction of Horizon 2020 funding to the non-occupied territories would have little impact on Israeli scientists, almost all of whom work within pre-1967 borders. Only one beneficiary of FP7 grants, cosmetics company Ahava Dead Sea Laboratories in Mitzpe Shalem in the West Bank, would have been ineligible had the guidelines been in place for FP7, and currently only about €1.5 million flows into the occupied territories. Both sides are considering how the guidelines could be implemented in a mutually acceptable manner, and will meet again in a few weeks, Jennings said.

Ruth Arnon, president of the Israel Academy of Sciences and Humanities in Jerusalem, wrote to Prime Minister Benjamin Netanyahu last month, warning him that Israel's failure to join Horizon 2020 would damage the country's science irreversibly. Others have warned that it is up to the scientific community to save the programme. "Whatever their individual political views, Israeli scientists need to come together to shout their concerns about a future without Horizon 2020 — but they have shied away from this," says Yadin Dudai, a neuroscientist at the Weizmann Institute.

Dudai is especially concerned about Israel's future involvement in the Human Brain Project, an 80-partner EU flagship programme that aims to unlock the mysteries of the mind and develop treatments for brain diseases. The project is currently in a ramp-up phase that is being partially funded by FP7; eight of its principal investigators are Israeli, including Dudai.

Gal Richter-Levin, a neuroscientist at the University of Haifa, says it is good for scientists to rub shoulders with their EU counterparts. "When we collaborate with colleagues in other countries, we have the opportunity to be criticized by them over a beer in the evenings," he says. "This helps us to form our political views — and this is also a reason for not abandoning scientific collaborations." ■



FRANK KRAHMER/GETTY

Possible infection with the MERS coronavirus, or a closely related virus, has been detected in camels.

#### VIROLOGY

# Progress stalled on coronavirus

*Lack of in-depth studies hampers efforts to identify source.*

BY DECLAN BUTLER

A year on from the first reported human case of infection with Middle East respiratory syndrome coronavirus (MERS-CoV), the world still has few answers to the most pressing question from a public-health perspective: what is the source of the steady stream of new cases? Only with this information can the outbreak be controlled.

There have so far been 114 confirmed cases of MERS-CoV infection, including 54 deaths, with another 34 suspected cases (see 'Catching on'). All originated in the Arabian Peninsula, with most in Saudi Arabia and others in Jordan, Qatar and the United Arab Emirates (UAE). Imported cases have occurred in the United Kingdom, France, Italy and Tunisia. The virus is thought to be an animal virus that sporadically jumps to people — there are no signs yet that it can spread easily between humans, although limited spread between people in close contact has been seen.

But a lack of epidemiological studies means that the source of the virus is still unknown. This is "absolutely unacceptable", says Michael Osterholm, head of the University of Minnesota's Center for Infectious Disease Research and Policy in Minneapolis. Affected countries have to some extent lacked a sense of urgency, agrees Jean-Claude Manuguerra, head of the Laboratory for Urgent Response to Biological Threats

at the Pasteur Institute in Paris. They have also tended to consider it a sovereign issue that is the business of national authorities, not outsiders, he adds.

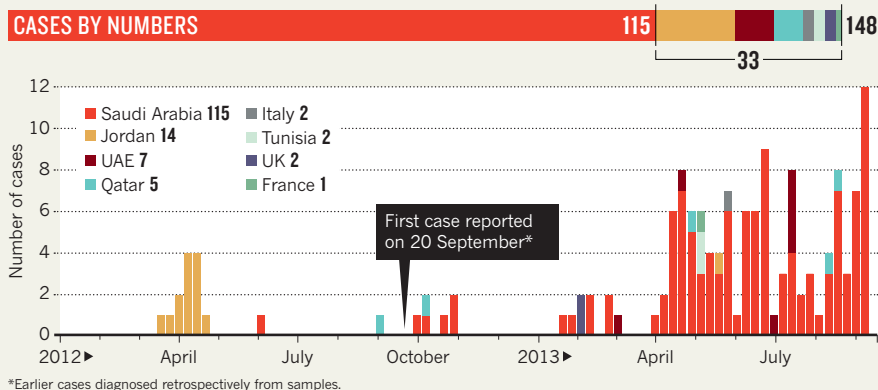
Manuguerra was one of 13 international health experts invited by the Saudi Arabian government and the World Health Organization (WHO) in Geneva, Switzerland, to visit the country in June as part of a joint mission to assess the situation and recommend measures to control the outbreak. He contrasts the lack of progress on MERS epidemiology with China's rapid and robust response following the report in March of the first human cases of a new H7N9 avian influenza virus (see *Nature* **496**, 397; 2013), which quickly identified live-bird markets as the main route of human infections and moved swiftly to shut them down.

To identify a source of infection, patients must be questioned in detail about their recent activities, such as whether they visited animal markets or farms, or ate certain foods, to find patterns and leads for investigation. But the joint Saudi-WHO mission concluded that such work has not been detailed enough to accomplish this.

Also insufficient, the mission said, were efforts to look for the virus in animal populations, which would enable transmission routes to humans to be shut down. Bats are likely to have been the original source: the genome sequence of the MERS virus is closest to that of

## CATCHING ON

There have been a total of 148 cases of MERS, although not all of these have been confirmed by the WHO.



more, he adds, and the FAO hopes to work with national authorities to start epidemiological studies within months.

A tantalizing clue to a possible animal source comes from two studies on camels<sup>2,3</sup>. Both found antibodies to MERS or a closely related virus in camels in neighbouring countries to those affected, indicating past infection. Neither found live virus in the animals, however.

“It is possible that such a virus is ubiquitous, but from time to time is undergoing crucial mutations that allow limited cross-species jumps from the intermediate host to humans,” says Malik Peiris, a virologist at the University of Hong Kong and a co-author of one of the studies. That was the scenario with the exotic coronavirus SARS that emerged in Hong Kong in late 2002 and evolved to become a virus that spread easily between humans, killing almost 800 people before dying out in July 2003.

With millions soon to be on the move for the hajj, the annual pilgrimage to Mecca in mid-October, the Saudi authorities are on high alert for suspected cases of MERS, and will no doubt be hoping that MERS is not the next SARS. ■

1. Memish, Z. A. *et al. Emerg. Infect. Dis.* <http://dx.doi.org/10.3201/eid1911.131172> (2013).
2. Reusken, C. B. E. M. *et al. Lancet Infect. Dis.* [http://dx.doi.org/10.1016/S1473-3099\(13\)70164-6](http://dx.doi.org/10.1016/S1473-3099(13)70164-6) (2013).
3. Perera, R. A. *et al. Eurosurveillance* **18**(36), 2 (2013).

other bat coronaviruses. Since this discovery, Saudi Arabia has twice invited a team led by Ian Lipkin, a virologist at Columbia University in New York, to sample bat tissue and faeces, and in August the team published a partial sequence that seems to match that of MERS<sup>1</sup> (see *Nature* <http://doi.org/ntk; 2013>).

But direct contact with bats is unlikely to explain the human cases, and scientists suspect that the virus has infected other animals that are more likely to come into contact with people. Some say they are baffled that affected

countries have not carried out large-scale sampling of animals to look for the virus. Ziad Memish, Saudi Arabia's deputy minister of health, says that his country has been waiting for advice about testing from the Rome-based United Nations Food and Agriculture Organization (FAO) and the World Organisation for Animal Health in Paris.

Juan Lubroth, the FAO's chief veterinary officer, says that its first expert mission to assess possible animal sources left last weekend for Saudi Arabia and the UAE. There will be several



# OUTLOOK FOR EARTH

## AS THE IPCC FINALIZES ITS NEXT BIG CLIMATE-SCIENCE ASSESSMENT, *NATURE* LOOKS AT THE PAST AND FUTURE OF THE PLANET'S WATCHDOG.

In December 1988, the United Nations General Assembly endorsed a call to create a panel to assess “the magnitude, timing and potential environmental and socio-economic impact of climate change and realistic response strategies”. Now in its 25th year, the resulting Intergovernmental Panel on Climate Change (IPCC) has grown substantially from its early days, when just a few dozen experts convened to write its first scientific assessment report.

Next week, the group will publish its fifth such report, which has been crafted by more than 250 lead authors and editors — as well as hundreds more contributors and reviewers — who spent five years on the project and had to deal with a flood of some 52,000 comments submitted in response to early drafts (see page 298).

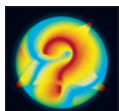
As the IPCC has matured, it has become firmer in its message that humanity is warming the globe to a degree that will threaten much of the world's population over the next century. The panel's warning has grown more confident and specific with time, as have its assessments of strategies to mitigate future problems.

In a special issue this week, *Nature* examines how the IPCC and climate science have evolved over the past quarter-century, and how scientific assessments can intersect with policy decisions. One News

Feature on page 300 investigates how much progress the panel has made on the topic of rising sea levels — one of the most controversial aspects of its previous big report, published in 2007. Another News Feature (see page 303) profiles economist Ottmar Edenhofer, who is leading the IPCC working group that will next April issue a second report, on ways to mitigate climate change and its impacts. A third IPCC report, on the impacts of climate change, will come out next March.

Even as the research on global warming has progressed steadily, nations have failed to forestall the problem. In a Comment on page 307, climate-policy analyst Elliot Diringer argues that individual actions by countries may be the best way forward. And on page 310, energy analyst K. John Holmes offers a look at how large-scale environmental assessments informed policy debates in the nineteenth century.

The IPCC has won praise for its work, including a Nobel Peace Prize in 2007, yet many researchers question whether it should continue to produce giant, infrequent reports. Past assessments have already described the basic science on climate change, which is now broadly accepted by governments. The IPCC can still play an essential part in climate mitigation, but it is time to re-evaluate how it offers advice to nations. ■ [SEE EDITORIAL P.281](#)



### OUTLOOK FOR EARTH

A *Nature* special issue on the IPCC  
[nature.com/ipcc2013](http://nature.com/ipcc2013)

# 25 years of the IPCC

THE INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE (IPCC) was founded 25 years ago to provide authoritative assessments on the emerging problem of climate change. Since its first report in 1990, the IPCC has issued increasingly complex follow-ups about every six years. The climate models that feed into the assessments have grown bigger and better, but researchers have not succeeded in reducing some key uncertainties about climate change. Where the reports have grown most firm is in declaring that humans are causing the world to warm.

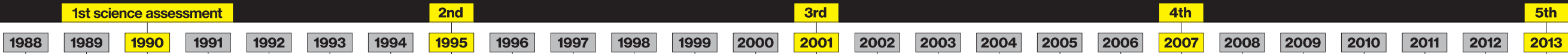
**GETTING BIGGER**  
The IPCC's Working Group I report on the science of climate change has more than quadrupled in length. The newest, due later in September, drew 52,822 comments in response to its first and second drafts. Source: IPCC

**INCREASING COMPLEXITY**  
The models for forecasting climate change have grown more detailed and now consider ocean biology, soil processes and atmospheric chemistry. Their resolution has improved and the number of models has increased.

**UNEVEN PROGRESS**  
Researchers have made no gains in narrowing the estimate of 'climate sensitivity': how much the world is likely to warm if atmospheric CO<sub>2</sub> levels double from the preindustrial value of 270 parts per million (p.p.m.). But the models have become more accurate at simulating the current climate. Source: IPCC

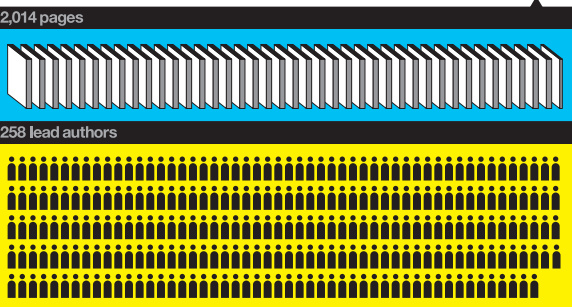
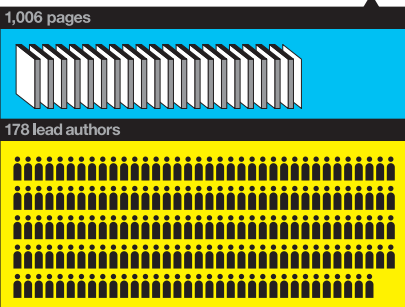
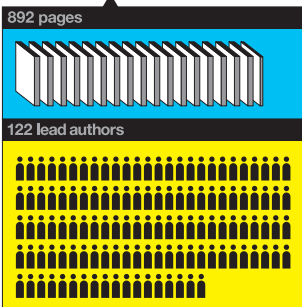
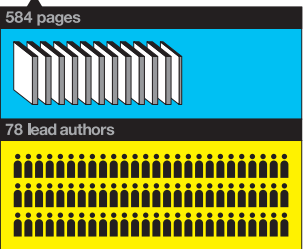
**GAINING CONFIDENCE**  
Successive IPCC reports have grown more definitive in identifying humans as the cause of much of the recent warming. Source: IPCC

Annual average CO<sub>2</sub> levels were about 350 p.p.m. in 1988. Source: NOAA

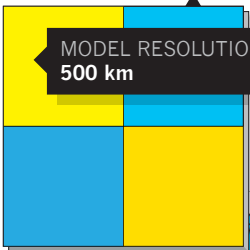


**REPORT PAGES**  
Each paper ream represents 50 pages.

**LEAD AUTHORS** Count does not include contributing authors, who often number many hundreds more.

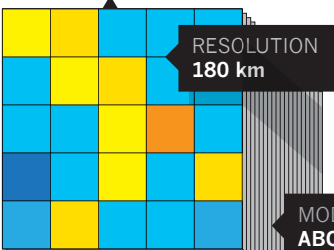
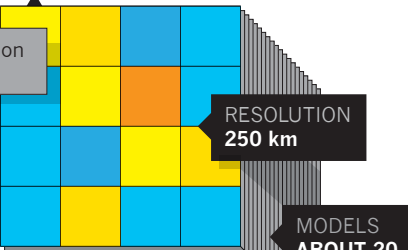


Average CO<sub>2</sub> levels will pass 400 p.p.m. in the next few years

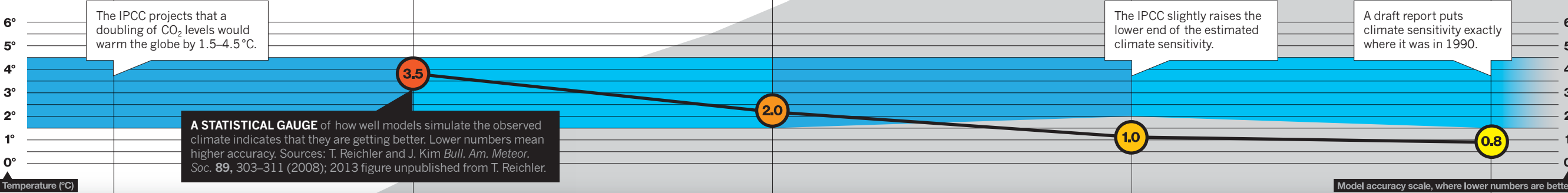
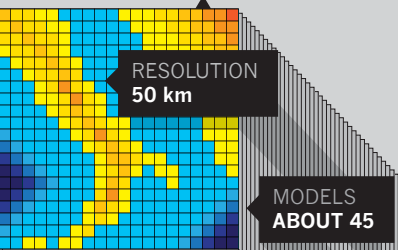
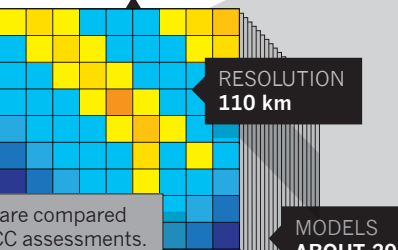


The most detailed horizontal grid resolution used for short-term climate simulations.

MODELS AVAILABLE ABOUT 2



Different models are compared as part of the IPCC assessments.



IPCC established.

**1ST ASSESSMENT REPORT**  
"Emissions resulting from human activities are substantially increasing the atmospheric concentrations of the greenhouse gases... These increases will enhance the greenhouse effect, resulting on average in an additional warming."

**EARTH SUMMIT** paves the way for future treaties limiting greenhouse-gas emissions.

**2ND ASSESSMENT REPORT**  
"The balance of evidence suggests a discernible human influence on global climate."

**KYOTO PROTOCOL SIGNED**  
The first, and so far only, multinational pact to reduce greenhouse-gas emissions.

**3RD ASSESSMENT REPORT**  
"Most of the observed warming over the last 50 years is likely [>66% chance] to have been due to the increase in greenhouse gas concentrations."

**KYOTO PROTOCOL** comes into effect.

IPCC shares the Nobel Peace Prize.

**4TH ASSESSMENT REPORT**  
"Most of the observed increase in global average temperatures since the mid-20th century is very likely [> 90% chance] due to the observed increase in anthropogenic greenhouse gas concentrations."

**KYOTO PROTOCOL EXPIRES**  
Participating nations cut emissions by 16%, but global emissions are still up 50% over 1990 levels.

**5TH ASSESSMENT (LEAKED DRAFT)**  
"It is extremely likely [>95% chance] that human influence on climate caused more than half of the observed increase in global average surface temperature from 1951–2010."





# RISING TIDE

RESEARCHERS  
STRUGGLE TO PROJECT  
HOW FAST, HOW HIGH  
AND HOW FAR THE  
OCEANS WILL RISE.

BY NICOLA JONES

The world's leading climate scientists kicked up a storm in 2007, when they issued their best estimates of how quickly the oceans would swell as the globe warms. The Intergovernmental Panel on Climate Change (IPCC) projected that sea levels would rise by somewhere between 18 and 59 centimetres by the last decade of this century — an upper limit that seemed far too low to other scientists, given the pace of melting in Greenland and other changes. “We were hugely criticized for being too conservative,” says Jerry Meehl, a climate modeller at the US National Center for Atmospheric Research in Boulder, Colorado and one of the authors of the IPCC’s 2007 report<sup>1</sup>.

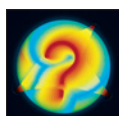
The panel had previously projected much higher rates of sea-level rise, but its 2007 assessment admitted that it could not tackle the entire problem: the predictions did not include the possibility of rapid changes in ice cover in Greenland or the Antarctic because the authors had concluded that it was impossible to forecast such behaviour with the knowledge and models then available. Yet as early as 2009, it was clear that real sea-level rise was on pace to exceed the 2007 projections<sup>2</sup>.

As the IPCC prepares to release its latest summary of climate science next week, researchers say that they now have a better grasp of the problem. Although the final report is not yet complete and the numbers could change, a leaked draft from June forecast a significantly greater rise in sea level — possibly close to 1 metre by 2100. But there is still huge uncertainty over how fast the oceans will rise, how the pattern will vary around the globe and what the ultimate high-water mark will be. Here, *Nature* investigates some of the big questions remaining about sea-level rise.

## HOW FAST WILL IT RISE?

Stefan Rahmstorf, a physical oceanographer at the Potsdam Institute for Climate Impact Research in Germany, is deeply unsatisfied with the standard tools for forecasting sea-level rise: ‘process’ models that try to represent the physics of every contributing factor. One reason for this discomfort was clear back in 2007. When researchers added up all the individual processes that contributed to rising seas, they could account for only 60% of the observed lift from 1961 to 2003 (see ‘Too Much Water’). “The whole was bigger than the sum of its parts,” says John Church, co-lead author of the chapter on sea-level rise in the forthcoming IPCC report and an oceanographer at the Australian Commonwealth Scientific and Industrial Research

PETER RYAN/NATL GEOGRAPHIC/GETTY



**OUTLOOK FOR EARTH**

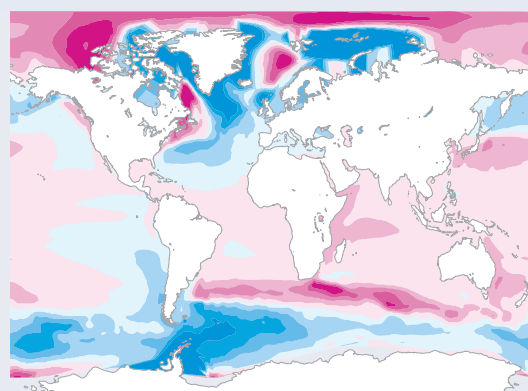
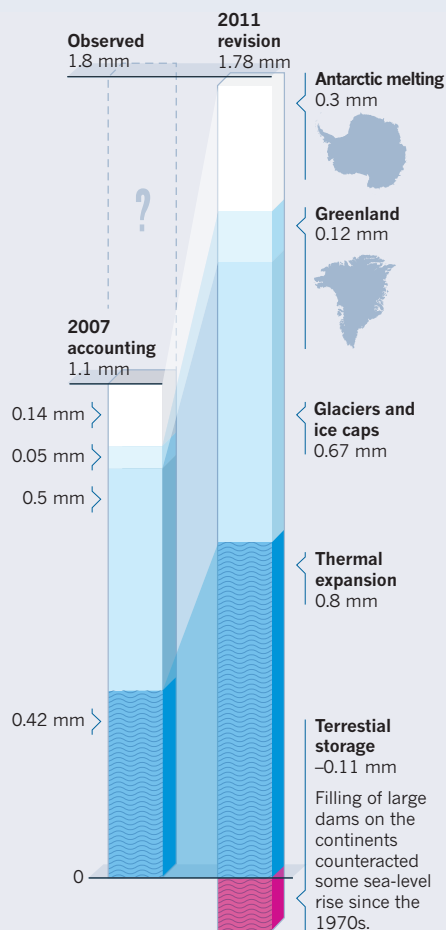
A *Nature* special issue on the IPCC  
[nature.com/ipcc2013](http://nature.com/ipcc2013)

# TOO MUCH WATER

Sea levels have been edging upwards at an accelerated rate over the past decade, and the pace will quicken as the world warms.

## BETTER BUDGET

In 2007, researchers could account for only about 60% of the changes in sea level since the 1960s, during a period when the oceans were rising an average of 1.8 millimetres per year. But thanks to improved models, oceanographers can now explain almost all of the changes since the 1970s. Over the past two decades, the rate of rise has picked up, to more than 3 millimetres per year.



## HOT SPOTS

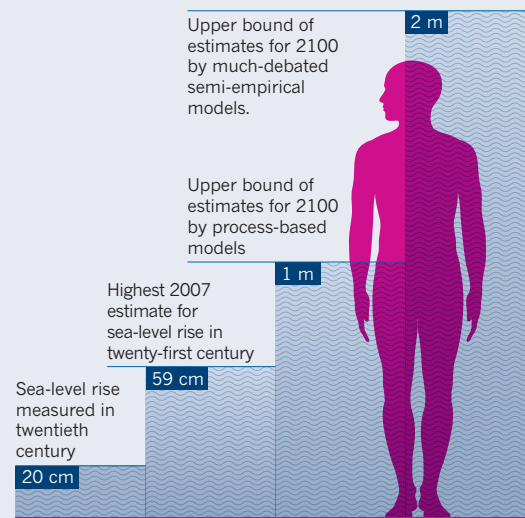
For a scenario with an average global sea-level rise of 47 centimetres, local factors will cause the oceans to rise more in some regions and less in others — and even drop in some.

Differences from global average



# 187 million

... people would be swamped by a 2-metre rise in sea levels. That is roughly the combined population of the United Kingdom, Germany and Spain.



SOURCE: 2007 BUDGET: IPCC; 2011 BUDGET: J.A. CHURCH ET AL. GEOPHYS. RES. LETT. 38, L18601 (2011); MAP: REF. 7

Organisation in Hobart. The two biggest effects — the expansion of water as it warms, and the addition of water to the oceans from melting glaciers — each accounted for about one-quarter of the total. A little extra was added in from the melting of the Antarctic and Greenland ice sheets. That left a gaping hole.

So Rahmstorf decided to pursue an entirely different type of model. He looked at the annual rate of sea-level rise from the 1880s onwards, and then matched it with air temperatures at those times. He found a simple relationship: the warmer it got, the faster the sea level rose. In 2007, too late to be considered by that year's IPCC assessment, his model predicted<sup>3</sup> up to 1.4 metres of sea-level rise by 2100 — more than twice the IPCC number.

'Semi-empirical' models such as this have advantages: by definition, they accurately model the rise that has already occurred, and they do not require a full understanding of how and why it is happening. But no one knows how long the relationship at the heart of these models will hold, particularly as melting ice sheets become a bigger factor. The models, says Rahmstorf, "could be good for 50 years, or 100 years. We don't know."

When it comes to making projections, the choice of models has big consequences. Process models generally predict rather less than 1 metre of rise by 2100, whereas semi-empirical

models top out at between 1 and 2 metres — enough, at the higher end, to flood the homes of 187 million people. These high-end, semi-empirical estimates are extremely controversial, and the IPCC has low confidence in them. "The only advantage of these models is that they're easy to calculate," says Philippe Huybrechts, an ice modeller at the Brussels Free University. "I think they're wrong."

Process-based modellers have made great progress since 2007, thanks to improved understanding of factors such as how much heat is flowing into the oceans — and thus causing the water to expand — and how much groundwater makes its way into the oceans because of people's unquenchable thirst for fresh water pumped up from below. As a result, modellers can now explain all of the observed rise in sea level, particularly that in recent decades.

But that does not guarantee accurate forecasts. Everyone acknowledges that there are still big issues with process-based projections — in particular, modellers have only a tenuous grasp of how the big ice sheets in Greenland and, especially, the Antarctic might behave and whether they will melt and flow catastrophically into the sea. In all, the ice sheets hold enough water to raise sea levels by more than 65 metres in the long term, compared with as much as 0.4 metres from all the world's glaciers and ice caps.

Despite these problems, the IPCC has

decided that researchers finally have a good enough handle on ice behaviour in Greenland and — to a lesser extent — Antarctica to forecast how ice sheets will respond, at least provisionally, says Don Chambers, a sea-level researcher at the University of Texas at Austin. The latest estimates add between 3 and 21 centimetres to the predicted sea-level rise by 2100, although tens of centimetres more are possible, according to the most recent IPCC report draft.

The end result is set to be a much higher forecast for sea-level rise than in 2007. Direct comparisons are difficult because the latest report uses different time frames and emission scenarios, but the leaked draft puts the range of estimates between 28 and 97 centimetres of rise by 2100. That is still not as high as semi-empirical estimates, but process-based results are edging upwards — and the difference is narrowing. "I consider it something of a vindication," says Rahmstorf.

## HOW MUCH WILL IT VARY?

When Jeff Freymueller, a geophysicist at the University of Alaska Fairbanks, visited Alaska's Graves Harbor more than a decade ago, his marine charts showed three isolated little islands; what he saw, instead, were three grassy peninsulas connected to the mainland. That was because water levels in some parts of Alaska are dropping — by up to 3 centimetres per year.



The ground there is lifting upwards, in a slow-motion rebound that has been going on for 10,000 years, since the glacial ice sheet that once weighed down the continent receded at the end of the last ice age. Gravitational influences on the oceans are also at work: as local glaciers recede and the Greenland ice sheet melts, their gravitational pull is subtly reduced, allowing more ocean water to slop southwards.

Trends in local sea level can differ strongly from the global average, which is increasing by around 3.2 millimetres per year. "Some places, sea-level rise is ten times faster than the average," says Jerry Mitrovica, a geophysicist at Harvard University in Cambridge, Massachusetts.

One side of this equation is the movement of the land. Canada's Hudson Bay, for example, was once buried under more than 3 kilometres of ice, and the release from that load is now causing the land to rise at about 1 centimetre per year. As that part of North America moves upwards, land to the south is being levered down: the US east coast is dropping by millimetres per year.

Subsidence can cause some areas to sink much faster. Compaction of river sediments and hollowing out of the earth by groundwater extraction, for example, are causing parts of China's Yellow River delta to sink at up to 25 centimetres per year<sup>4</sup>.

Adding to the complexity, the oceans do not rise evenly all over the world as water is poured in. Air pressure, winds and currents can shove water in a given ocean to one side: since 1950, for example, a 1,000-kilometre stretch of the US Atlantic coast north of Cape Hatteras in North Carolina has seen the sea rise at 3–4 times the global average rate<sup>5</sup>. In large part, this is because the Gulf Stream and the North Atlantic current, which normally push waters away from that coast, have been weakening, allowing water to slop back onto US shores.

Finally, waters near big chunks of land and ice are literally pulled up onto shores by gravity. As ice sheets melt, the gravitational field weakens and alters the sea level. If Greenland melted enough to raise global seas by an average of 1 metre, for example, the gravitational effect would lower water levels near Greenland by 2.5 metres and raise them by as much as 1.3 metres far away.

Scientists and engineers are only just starting to wrangle all these effects into local projections. In June, the New York City Panel on Climate Change updated its estimates of sea-level rise by including the local effects of gravitational shifts<sup>6</sup>. Panel members concluded that they expect to see 30–60 centimetres of rise by 2050. Finding and combining the right data sets took about six months; the exercise should pave the way for other cities to do the same, says Cynthia Rosenzweig, a climate-impact researcher at NASA's Goddard Institute for Space Studies in New York City. "We really are working to get the best science."

Aimée Slangen, now a member of Church's

group in Australia, last year put out one of the first global maps of regional sea-level change that takes all these factors into account, but it had only rough resolution, with pixels of more than 100 kilometres each<sup>7</sup>. Researchers want to provide city-level predictions, but are hampered because these depend heavily on decadal shifts in winds and ocean currents. Predicting such changes is "very problematic", says Chambers.

Regional figures for sea-level rise are of interest not only to people trying to plan for local impacts, but also to those trying to model global effects. For the latter, the numbers bring some good news. Gravitational shifts caused by melting ice in the Antarctic should actually help to prevent catastrophic collapse of the West Antarctic ice sheet: as Antarctica loses some of its ice, local sea levels will fall,

**"SOME PLACES, SEA-LEVEL RISE IS TEN TIMES FASTER THAN THE AVERAGE."**

which will cause some floating edges of the ice sheet to come to rest on the sea floor. Firmly grounded ice is less susceptible to runaway melting than floating ice. "That's going to stabilize the ice sheet," says Mitrovica.

#### HOW HIGH WILL IT GO?

"Sea-level rise isn't going to stop in 2100," says Church. "I think that's something that people don't really take on board." Eventually, they will. Projections of sea-level rise far into the future jump from tens of centimetres to tens of metres.

For the past few years, Maureen Raymo, a marine geologist at the Lamont-Doherty Earth Observatory in Palisades, New York, has traipsed around abandoned diamond mines in South Africa, visited quarries in Australia and examined road cuts on the east coast of North America, looking for shells and other remnants of beaches from 3 million years ago. She hopes to reconstruct sea levels from the Pliocene epoch, the last time when carbon dioxide concentrations were as high as they are today: about 400 parts per million of the atmospheric volume. That, in turn, should provide a glimpse of what the world might look like in thousands of years, once the planet has had time to react fully to today's emissions.

Current estimates of sea-level rise in the Pliocene range from very little to 40 metres, says Raymo. "But that's not very helpful," she says. The difference between the lower and higher estimates is the difference — crucially — between much of the vast East Antarctic ice sheet melting and staying frozen. Whether or not it melted in the Pliocene, in turn, provides insight for modellers who are trying to work out whether — and how fast — ice sheets might collapse in the next few hundred years.

The trick to pinning this down is not just finding Pliocene beaches, but also working out

how the land has moved since they were laid down, as a result of both rebound from the loss of ice sheets and the ongoing movement of mantle rock under the continents. To estimate how such processes have played out over millions of years, researchers rely on models of how much ice covered the continents and how viscous the mantle is — factors that are subject to extreme debate. "Today's models all assume a viscosity of the mantle that is untestable, highly controversial and differs between groups," says Raymo.

This motion of Earth makes a big difference in estimating past events. Previous work, for example, in Bermuda and the Bahamas hinted that the coastline there was 20 metres higher than it is today during a warm period 400,000 years ago. In 2012, however, Raymo and Mitrovica calculated how the ground there had moved — and concluded<sup>8</sup> that half of the apparent sea-level rise was attributable not to rising waters but to sinking land, cutting the sea-level-rise estimate in half.

Given the large error bars, the only way to pinpoint Pliocene sea levels is to get data from many sites and to calculate one best-fit answer for global sea level. Raymo and her team have so far surveyed thousands of kilometres of coast, gathering evidence from dozens of beach sites. She says she needs perhaps eight more locations and five years to finish the job.

But, she admits, whatever she finds will not be a worst-case scenario because greenhouse-gas concentrations are already climbing beyond where they were in the Pliocene. "The real worst-case scenario is we don't limit fossil-fuel combustion," she says. "Then it's 'Hello Eocene'" — returning to a world akin to a warm period 55 million years ago, with maybe just a trace of ice at the poles.

Nearly 70 metres of sea-level rise would drown all of Florida and much of Brazil, and swamp the Statue of Liberty up to her waist. But that might not happen until so many thousands of years from now that humanity has time to adapt — even if that means surrendering much of the land to the waves. ■

**Nicola Jones** is a freelance writer near Vancouver, Canada.

1. Solomon, S. et al. (eds) *Climate Change 2007: The Physical Science Basis* (Cambridge Univ. Press, 2007).
2. Allison, I. et al. *The Copenhagen Diagnosis 2009* (Univ. New South Wales Climate Change Research Centre, 2009).
3. Rahmstorf, S. *Science* **315**, 368–370 (2007).
4. Higgins, S., Overeem, I., Tanaka, A. & Syvitski, J. P. M. *Geophys. Res. Lett.* **40**, 3898–3902 (2013).
5. Sallenger, A. H. Jr, Doran, K. S. & Howd, P. A. *Nature Clim. Change* **2**, 884–888 (2012).
6. New York City Panel on Climate Change *Climate Risk Information 2013* (City of New York Special Initiative on Rebuilding and Resiliency, 2013).
7. Slangen, A. B. A., Katsman, C. A., van de Wal, R. S. W., Vermeersen, L. L. A. & Riva, R. E. M. *Clim. Dynam.* **38**, 1191–1209 (2012).
8. Raymo, M. E. & Mitrovica, J. X. *Nature* **483**, 453–456 (2012).



# THE CLIMATE CHAIRMAN



**GETTING HUNDREDS OF EXPERTS TO AGREE IS NEVER EASY. OTTMAR EDENHOFER TAKES A FIRM, PHILOSOPHICAL APPROACH TO THE TASK.**

BY QUIRIN SCHIERMEIER

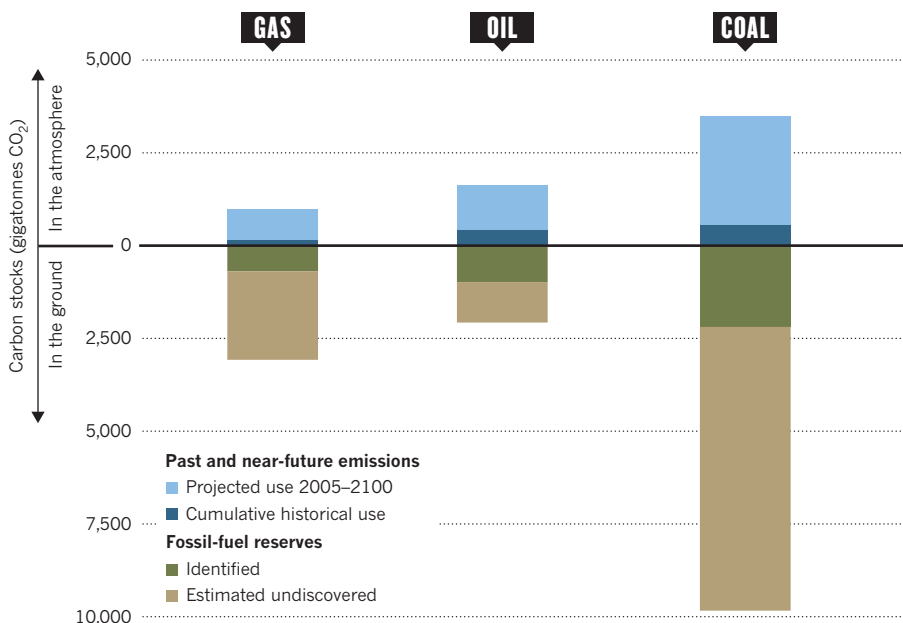
Ottmar Edenhofer knows that he sometimes has trouble keeping his composure. So when he takes the front seat as chairman at international climate meetings, he attaches an index card to his desk and glances at it whenever the discussions turn testy. The sign warns: "Don't lose your temper!"

He has needed that reminder often during his five years as one of the leaders of the Intergovernmental Panel on Climate Change (IPCC), where science frequently comes into conflict with powerful political interests as nations debate how climate is changing and what we can do about it. Those tensions nearly boiled over in 2011, when Edenhofer presided over a session in Abu Dhabi at which delegates from some 200 nations fought over almost ►



## THE CARBON AGE

The amount of carbon dioxide that humans have pumped into the atmosphere so far pales in comparison to the volume that would be produced if all remaining reserves were burned.



► every single sentence in the summary of a report on renewable energies.

"It's an emotional roller coaster," says Edenhofer, chief economist at the Potsdam Institute for Climate Impact Research in Germany. "Things can get pretty nasty when delegations play their little games of power, and the chairs must be prepared to parry attacks. My little reminder to always stay cool and polite has saved me from fits more than once."

His temper will be tested again during the coming months as he guides his committee — the IPCC's Working Group III, which looks at ways to mitigate climate change — towards the end of its five-year reporting process. A week-long meeting in late June and early July in Addis Ababa turned into a marathon for the 232 lead authors of the report, who were faced with addressing more than 16,000 comments submitted by expert reviewers and governments in response to an early draft. Next month, Edenhofer's committee must send a revised version to governments to prepare for April 2014, when lead authors and government representatives will meet in Berlin to hammer out the final report.

## BLUEPRINTS FOR A GREEN FUTURE

Edenhofer will need to marshal a unique blend of skills, honed during a stint as a Jesuit philosophy scholar and through research on game-theory, as he completes the most sweeping compendium yet of technology and policy options that might spare humanity from the worst of climate change in the coming decades.

The report will lay out a range of scenarios — and the costs and risks of each — for transforming societies to stabilize greenhouse-gas concentrations at reasonably safe levels.

The results will inform the political process through a round of global climate negotiations intended to culminate in a treaty in 2015. And as nations continue to disagree over the architecture and ambitions of that pact, the IPCC's take on mitigation will come under heavy scrutiny. The IPCC is charged only with laying out the science — and, in the past, critics have accused the group's leaders of overstepping the boundary between analysing research and advocating for action.

Well aware of that risk, Edenhofer has strived to keep his group focused on a strictly scholarly agenda — and, in so doing, he has won over some of the IPCC's past critics.

**"THINGS CAN GET PRETTY  
NASTY WHEN DELEGATIONS  
PLAY THEIR LITTLE  
GAMES OF POWER."**

"Ottmar has amazing skills," says Robert Stavins, an environmental economist at Harvard University in Cambridge, Massachusetts, who did not work on the last IPCC report out of concerns that it had grown too political. "I can't think of anybody better fitted for the job."

Philosophy was Edenhofer's intellectual refuge early on. Hailing from an

arch-conservative part of Bavaria in south Germany, he shocked his parents — and prompted the local book seller to declare him crazy — when he spent his savings at the age of 14 on the three volumes of *Das Kapital* by Karl Marx. Reading the German philosopher's critique of the capitalist mode of production didn't turn him into an ardent Marxist, but did spark his interest in political economics. A couple of years later, the writings of German sociologist Max Weber prompted him to ponder the value of science and the difficult relationship between values and facts — problems that are right at the core of any serious debate on the role of the IPCC.

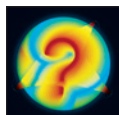
In the mid-1980s, Edenhofer studied economics at the University of Munich, Germany, but his academic career soon took an unusual detour. In 1987, he joined the Jesuit order in Munich to immerse himself in Western philosophy and, later, in theology. Soaking up the works of Weber, Ludwig Wittgenstein and John Dewey, he learned to embrace different lines of reasoning. In 1991, he also became involved in setting up a Jesuit Refugee Service in Bosnia and Croatia during the Yugoslav Wars, and later earned a PhD in economics at the Technical University of Darmstadt, Germany.

His roots in philosophy are still palpable in the way that he approaches the climate conundrum. "There is a whole space of morally legitimate standpoints with a view to climate change," he says one day in May, while sipping coffee between sessions of a workshop on climate agreements in Berlin. His boyish face, framed by round spectacles, grows animated as he lays out the various perspectives. "One might legitimately argue that the fight against global warming is as morally imperative as abolishing child labour or slavery. One might argue — just as legitimately — that poverty and diseases in many parts of the world are more imminent problems that should be addressed first."

However, he adds, some perspectives cannot be tolerated. "Denying out-and-out that climate change is a problem to humanity, as some cynics do, is an unethical, unacceptable position."

The upcoming Working Group III report that Edenhofer is presiding over is a massive, complex tome. A compendium of hundreds of scientific papers, it analyses how societies can slow down climate change and reduce its effects by altering all sectors of the economy, from electricity production to transportation to building design. The importance of his group's work has grown as the greenhouse problem has worsened with no political solution in sight. Since the 2007 IPCC report, emissions of heat-trapping gases have continued to grow despite a global economic recession. In 2012, annual emissions of such gases were equivalent to more than 50 billion tonnes of carbon dioxide and reached an all-time high. And, in May 2013, the concentration of atmospheric CO<sub>2</sub>

SOURCE: IPCC



**OUTLOOK FOR EARTH**

A Nature special issue on the IPCC  
[nature.com/ipcc2013](http://nature.com/ipcc2013)

crossed the ominous threshold of 400 parts per million (p.p.m.) for the first time since human beings appeared on Earth.

The report will provide a range of scenarios, and cost estimates, for stabilizing atmospheric CO<sub>2</sub> concentrations at 450 or 550 p.p.m.. It will make it clear that all realistic stabilization scenarios are decidedly at odds with current emissions trends. In fact, if one factors in methane, nitrous oxide and other warming gases that are governed by the 1997 Kyoto Protocol climate treaty, the combined concentration has already surpassed the equivalent of 450 p.p.m. of CO<sub>2</sub>. As a result, stabilization plans must allow for nations to temporarily overshoot a target before concentrations might start to subside.

The report will also make clear that the problem will only grow worse without action. Known hydrocarbon reserves still buried in the ground may contain up to four times as much carbon as has been released into the atmosphere since the onset of the Industrial Revolution (see 'The carbon age').

But in line with the IPCC's mandate — which requires the group to be policy-neutral — the assessment of Working Group III will avoid promoting certain mitigation options over others. Maintaining that distance will be a crucial test of Edenhofer's leadership.

"The IPCC is a scientific body whose task is to compare and review the relevant literature," says Stavins. "Unfortunately, in my view, the IPCC has in the past overstretched its mandate. It has become too political — and that hasn't done it and the field any good." Stavins feels, for example, that prominent IPCC members transgressed by lobbying for green policies such as emission cuts and carbon taxes.

Edenhofer is keen to steer clear of such territory: he often compares the task of the IPCC to that of map-makers rather than to that of political advisers. But he also knows that the report, by necessity, will weigh in on politically charged issues such as nuclear power, biofuels and geoengineering. When the report's summary for policy-makers — its most-read and most-disputed section — goes up for debate in April, the fighting among diplomats and scientists will be even tougher than it was in past sessions. Together with his Cuban and Malian co-chairs, Edenhofer will have to see, as patiently as his nature allows, that the haggling over the tiniest words proceeds in a civilized and productive manner.

He faces the date — his last as an IPCC official — with a mixture of anxiety and gladiatorial anticipation. "Where I grew up, the boys were always ready for a fight," he says. "I'm not like that — but trust me, I do know when it's time to thump on the table."

The stakes are mounting on his group's report. With its analyses of low-carbon energy options and different policy paths towards stabilizing greenhouse-gas concentrations, the report will be an important contributor to the upcoming negotiations as nations try to

hammer out an international treaty over the next year.

When the last round of IPCC Assessment Reports came out in 2007, the body came under fire for including a few claims that lacked strong scientific support; in one high-profile gaffe, Working Group II uncritically repeated a baseless assertion that Himalayan glaciers would disappear by 2035. Edenhofer has pushed his own working group to thoroughly overhaul its procedures for evaluating scholarship to avoid such embarrassing blunders in his report. In particular, he has reduced the use of 'grey' literature — information not subjected to peer review, such as reports from environmental groups, governments and companies — to a minimum.

**"SCIENCE CANNOT,  
AND CAN'T BE EXPECTED  
TO, PROVIDE SIMPLE  
YES OR NO ANSWERS."**

And, perhaps drawing on his background as a philosopher and theologian, Edenhofer has also broadened the scope of the IPCC's mitigation working group to give ethical considerations greater weight. For the first time, professional philosophers have been invited to contribute to the report's opening chapters on equality, risks and sustainability issues, which set the scene for more technical sections.

But even as it details the various options that nations might take, the report will not take sides, he stresses — because that involves decisions based on values and priorities that fall to elected officials. "Science cannot, and can't be expected to, provide simple yes or no answers," he says.

To tread that neutral line, Edenhofer will have to control his own strong opinions, formed during his years as an economist and climate-policy expert. He favours cap-and-trade schemes, for example, over a direct carbon tax as the most effective way to cut emissions and promote climate-friendly innovation. And he is adamant that the global transformation to a low-carbon economy cannot be achieved — no matter what goals key emitters might commit themselves to — without substantially increasing the use of renewable energy in all sectors of the economy.

In spite of his own opinions, he has focused on delivering an unassailable product. "Ottmar is keen to get the best science for the next IPCC report, which he wants to lift to a new level of quality," says Massimo Tavoni, deputy coordinator of climate-change programmes at the Eni Enrico Mattei Foundation in Milan, Italy.

Edenhofer is also ready to acknowledge the limits of knowledge in his field. The most important uncertainty, he says, concerns the

reliability of economic models used to forecast the future. They rely on macroeconomic equations and assumptions that are often thwarted by real-world developments. Economists are well aware that although the models can anticipate broad trends, they have no ability to forecast disruptions such as major financial or political crises.

And great uncertainty remains over how nations will tackle climate change. How much will countries cooperate? To what degree will they rely on nuclear power? How quickly will renewable energy be deployed and at what price? Beyond those near-term concerns, researchers must also grapple with more distant potential mitigation strategies such as capturing and storing carbon on a massive scale, or large geoengineering projects aimed at rapidly staving off warming.

As Working Group III tackles such uncertainties, it will also wade into an increasingly contentious debate about the benefits of creating biofuels from plants and bacteria. Since the IPCC's 2007 report, a fast-growing body of literature has split over whether the indirect effects of growing crops for fuels do more harm than good to the climate.

Fears that excessive bioenergy production might cause food shortages make the debate even fiercer. "Clearly," says Edenhofer, "this is one of the most controversial issues we're dealing with. By now, debate over bioenergy has outstripped controversy over nuclear energy." The IPCC, he says, will summarize the pros and cons as authoritatively as scientific knowledge allows.

## TEARY FINALE

Edenhofer is convinced that the IPCC is better placed than any other group to address such thorny issues, because the final reports are vetted not only by scientists but also by political appointees from member nations. No purely science-led exercise could possibly have equal weight, he says. Although some scientists have started to question the utility of the IPCC, especially its drawn-out procedures, Edenhofer says the process should continue. The "miracle" of the IPCC, as he puts it, is that it forces governments to deal seriously with science.

"To be able to engage and criticize our work, governments do need to carefully read our reports," he says. "Here's a unique mechanism for bringing science to the very level of government leaders."

As he steers his group through that process, he will make frequent use of his desktop index card. But he has faith in the process. "When everything is said and done, and when even the most hard-boiled negotiators have tears in their eyes, it is the cause of science — and not power interests — that has the last word." ■

**Quirin Schiermeier** is a reporter for Nature in Munich, Germany.



# COMMENT

**HISTORY** Lessons from the first large-scale environmental assessments **p.310**



**FICTION** Thomas Pynchon's latest novel **p.312**

**ECONOMIES** Comparing accounts of innovation in Russia and the United States **p.313**

**MATHEMATICS** Martin Gardner's posthumously published autobiography **p.314**

ILLUSTRATION BY DENIS CARRIER



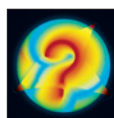
## A patchwork of emissions cuts

Home-made national approaches can be effective for climate-change mitigation if countries agree on rules and build trust, says **Elliot Diringer**.

**T**he last thing that most people recall about United Nations climate-change talks was the meltdown four years ago in Copenhagen, when presidents and prime ministers at the 15th Conference of the Parties (COP) to the UN Framework Convention on Climate Change (UNFCCC) failed to deliver the strong, binding treaty that the world had been led to expect.

Negotiations have, in fact, continued, and

governments have set another deadline: a new global climate agreement in 2015. With this latest round of talks nearing its midpoint, hints of a way forward are beginning to emerge. The path that they suggest



**OUTLOOK FOR EARTH**

A *Nature* special issue on the IPCC  
[nature.com/ipcc2013](http://nature.com/ipcc2013)

is at once encouraging and underwhelming.

Governments are yet to formally agree on anything. But in discussions earlier this year in Bonn, Germany, they began earnestly exploring a middle way — one neither as rigorous as the fading Kyoto Protocol to the UNFCCC, nor simply do-as-you-please. Although this approach might not look much like a grand climate solution, it is one that over time could substantially ►

► advance the global climate effort.

In the emerging model, which has been advanced by the United States among others, the efforts of countries would be tracked under agreed rules. But individual emissions-reduction goals would be set by each country on its own, without negotiation. The agreement would, in essence, stitch together a mixture of self-defined contributions. To encourage ambition, countries would scrutinize each other's initial offerings.

Having learned hard lessons in Copenhagen, governments are tempering their expectations this time around. Many now seem willing to accept a looser pact, if it can get everyone on board and offer the promise of strengthening action over time.

This thinking reflects in part the stubborn political reality that, despite the rising toll of extreme weather and other climate impacts, governments remain much more preoccupied with flagging economies than with climate disruption. It also reflects a sober reassessment of what the UNFCCC can effectively deliver. It is a concession to the limits of global diplomacy and law in tackling this truly global challenge, and a recognition that, whatever the UNFCCC's role is in moving forward, much of the real work to stave off climate catastrophe must happen at home.

## TWISTS AND TURNS

An early test for the emerging approach could come at COP19 this November in Warsaw. Some nations are likely to push for a formal process to begin eliciting and assessing countries' intended commitments. If the talks stay true to past form, however, definitive outcomes on the structure of an agreement, let alone the content, are unlikely to arise until the closing hours of COP21 in 2015.

This point in the UNFCCC's evolution may be understood by looking back to the framework's negotiation two decades ago. Competing governance models that dominated the debate then remain a source of tension. One model is the 'top-down' approach: countries agree on a long-term climate goal and a formula for allocating legally binding emissions-reduction targets that are stringent enough to achieve it. The other is a 'bottom-up' policy: countries offer up whatever commitments they like on a voluntary basis.

Elements of both models found their way into the UNFCCC, the parent agreement that has been ratified by 195 parties, under which the negotiations have proceeded since it entered into force in 1994. Parties leaned towards a top-down approach for the 1997 Kyoto Protocol, which set binding emissions targets for developed countries only, with the target levels negotiated rather than simply asserted.

Legally and technically, Kyoto introduced greater rigour into global climate governance. Politically, it has proven unsound. The

United States never ratified the protocol; Canada did, but later withdrew; and Japan and Russia refused to sign on for a second round of obligations. The current (and probably final) Kyoto targets, which run until 2020, cover only Europe and a handful of other countries — totalling less than 15% of global emissions. In the meantime, parties have spun together a parallel framework that takes a bottom-up approach. Copenhagen did produce a political accord that was formalized a year later in the Cancun Agreements, inviting voluntary emission pledges to be achieved by 2020.

More than 90 countries representing some 80% of global emissions — including, for the first time, all the world's major economies — submitted pledges under the agreements. But they are too weak to put the world on a pathway to keep warming below 2°C above pre-industrialized levels. The UN Environment Programme projects that, even in the best case, the pledges will achieve less than half of the reductions needed by 2020.

So governments have now experimented with both top-down and bottom-up approaches, and neither is getting the job done. Kyoto lacks adherents; the Cancun Agreements lack ambition. Is it possible to forge a middle path that has some chance of delivering both?

In launching the latest round of talks, parties agreed on little beyond a deadline (2015) and a start date for an agreement (2020). The mandate

**"Kyoto lacks adherents; the Cancun Agreements lack ambition."**

for the talks, part of a package deal struck at COP17 in 2011 in Durban, South Africa, requires "a protocol, another legal instrument or an agreed outcome with legal force under the Convention applicable to all Parties". An exemplar of constructive ambiguity, this hard-fought phrase manages to address, while leaving entirely unresolved, two perennial issues: legal form and the distribution of effort between developed and developing countries.

Since Durban, developing countries have continued to insist that wealthy nations bear a much greater responsibility for reducing emissions because, both cumulatively and per capita, they have generated so many more. Developed countries are equally adamant that the mandate ends the strict binary approach of the Kyoto Protocol: binding emissions targets for developed countries, and no commitments for developing countries.

The distinction between developed and developing countries would be less important in the type of agreement visualized by the United States. All countries' obligations would have the same legal character — if they are binding for some, then they are binding for all. But beyond that, countries would have

wide latitude in deciding the stringency and type of their emissions-reduction measures. The United States proposed in March that parties should "define their own mitigation contributions, taking into account national circumstances, capacity, and other factors that they consider relevant".

Most other parties seem willing to accept that the stringency of future commitments will be self-defined. But developing countries continue to call for differentiation in type, with developed countries adopting economy-wide emissions-reduction targets, and developing countries free to pursue other measures, as in the Cancun Agreements.

## BUILDING AMBITION

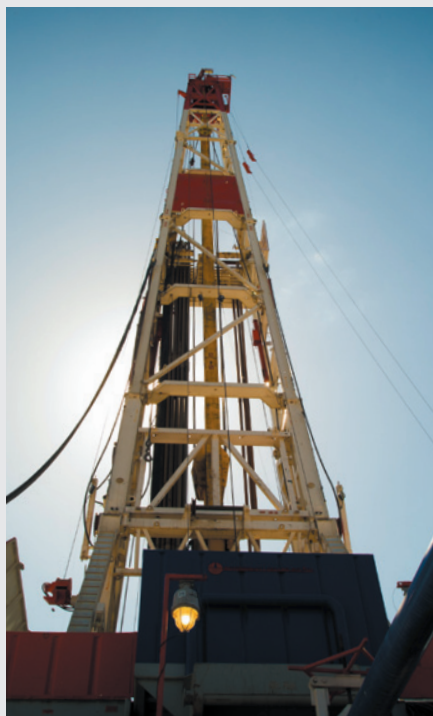
The major issue that negotiators are struggling with, however, is how to build ambition into the model, and how to ensure that collective action does reduce global emissions overall. Much of the thinking so far has centred on some form of 'ex-ante' review of commitments. If countries submit their intended offerings to scrutiny before they are formally inscribed, they might be more ambitious. A similar approach has been used for many years in the global trade regime to progressively reduce tariffs.

In the US view, the review would simply involve governments exchanging proposals and responses bilaterally. The European Union (EU) and others foresee a more open, formal process. In May, the EU outlined a 'stepwise approach' that would begin by spelling out the information needed from parties to ensure that their proposals are "transparent, quantifiable and comparable", and lead to a "collective consideration of [their] overall adequacy... and the increase of ambition if needed to stay on track for below 2°C".

Many parties also want a 'ratcheting' mechanism to boost ambition over time. Instead of a lengthy process of negotiation and ratification for each set of commitments, they foresee a dynamic, evolving agreement. Countries would report on how they are implementing their existing commitments, assess one another's efforts and then initiate another round of offers and reviews — a continuous cycle inducing a progressively stronger, collective effort.

If successful, such an arrangement could have an additive effect. Still, its fundamental force would derive less from the rule of international law than from the concrete national programmes that it would presumably stitch together. The most important quotient would remain domestic political will.

There are encouraging signs: China is pushing forward with carbon trading, and US President Barack Obama has promised to cut carbon emissions from US power plants. But at the same time, Germany and the United Kingdom are burning more coal, China's emissions continue to soar



Gas drilling in Pennsylvania; solar-panel installation in Jiuquan, China; a coal plant in Grevenbroich, Germany's 'energy capital'.

## CASE STUDIES

### Leading emitters

**United States.** Emissions are down from their 2007 peak, largely as a result of a slow economy, a natural-gas boom and advances in energy efficiency. Declaring climate change a top priority for his second term, President Barack Obama is promising to use his executive powers to reduce emissions further. Obama's directive to the US Environmental Protection Agency to regulate carbon from power plants could put the country in a position to meet the president's 2009 pledge at the Copenhagen climate summit of a 17% cut (from 2005 levels) by 2020. There remains little prospect of the politically fractured US Congress enacting anything more ambitious.

**China.** The country is forging ahead with the very market-based approach to cutting carbon that the US Congress rejected. Having seized a major share of the global renewable-technology market (upsetting US and European Union (EU) trade), China is piloting eight regional 'cap-and-trade' systems, with the goal of a national carbon market by 2020.

Yet China's dependence on high-carbon energy shows few signs of abating. The International Energy Agency in Paris projects that China's coal and oil consumption will rise by 23% and 46%, respectively, by 2020 — when its carbon dioxide emissions will have exceeded those of the United States and EU combined.

**Europe.** It has long been the most climate-committed of the three leading emitters, but rolling economic crises have weakened climate fervour. A push to strengthen the EU's 2020 emissions target (from 20% to 30% below 1990 levels) has been sidetracked by the recession.

Although sluggish economies in eastern Europe have kept overall emissions from the continent falling, coal use and emissions are edging up in Germany and the United Kingdom. Even as Germany invests heavily in renewable power, it is importing record levels of coal — a result of nuclear shutdowns following the 2011 Fukushima disaster in Japan and rising natural-gas prices.

and carbon pricing remains off the table in Washington DC (see 'Leading emitters').

The odds of reaching an agreement will be better in 2015 than they were in Copenhagen. With the Intergovernmental Panel on Climate Change scheduled to begin the staggered release of its Fifth Assessment Report next week, and costly climate impacts growing steadily, governments know that they must produce something. And they seem to be further along in understanding what that might be than they were last time.

The French government, which is slated to host the 2015 talks, is moderating expectations in the hope of avoiding a repeat of

Copenhagen. But whether or not anything firm is agreed in November, the parties will face mounting pressure to put preliminary targets on the table by the time UN secretary-general Ban Ki-moon gathers leaders for a climate summit in September 2014.

In the subsequent year, they will also need to decide how a new pact would address a host of other issues, such as helping poor countries to cope with climate impacts, the role of carbon markets and the agreement's legal force. Will countries' self-defined commitments be legally binding? Or perhaps only the periodic reporting of what they have achieved?

A new agreement will probably not be heroic, but it can be pragmatic. With creativity and compromise, governments can devise a multilateral framework that lends coherence, transparency and rigour to the emerging patchwork of national efforts. That, in turn, can strengthen countries' confidence in one another, in the process and in our collective ability to overcome the climate challenge. ■

**Elliot Diringer** is executive vice-president of the Center for Climate and Energy Solutions (C2ES) in Arlington, Virginia.  
e-mail: [diringere@c2es.org](mailto:diringere@c2es.org)





A party sets off from Wyoming on a scientific exploration of the western United States in May 1871.

# Pushing the climate frontier

The first large-scale environmental surveys, carried out on the US arid lands, hold scientific lessons for policy-making still relevant today, explains **K. John Holmes**.

When nineteenth-century explorer William Gilpin travelled across the Great Plains, the expanse that covers much of the central and western United States, he marvelled at the “great pastoral region”, the dry climate of which was “favorable to health, longevity, intellectual and physical development”<sup>1</sup>. Great cities could be built there, he imagined, taking advantage of the wealth of local resources — rivers, forests and even gold.

Geologist John Wesley Powell saw things differently. Moving from the humid east to the arid west would affect agricultural practices, occupations, social interactions and political customs, he contended<sup>2,3</sup>. Dry-land agriculture could not support a large population; any towns built in the west would need appropriate designs, irrigation and resource management. A controversy erupted.

The ensuing debates about how the arid lands should be settled hold lessons for us today on adapting to a changing climate. At their heart was a development plan for the region that Powell published in 1878 (ref. 2). It

called for detailed scientific and engineering surveys, and analysis to inform land-use plans and laws. Although it addressed a spatial change in conditions caused by westward population expansion, Powell’s coupling of physical and human dimensions was a forerunner to the assessment approach used today by the Intergovernmental Panel on Climate Change (IPCC).

Powell’s plan was never implemented in its entirety, but it began an era in which large-scale environmental and natural-resources assessments became central to the policy process in the United States<sup>4</sup>. Stalled by misinformation, political controversy and recessions, legislation for allocating resources in the arid lands took decades to enact. Then, as now, the assessments and their validity became part of the debate. Eventually, extreme weather, including long

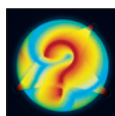
droughts, pushed policy-makers to act.

From New Mexico and the Dakotas to California’s Sierra Nevada mountains, the arid lands of the United States — dubbed the Great American Desert in the accounts of other early nineteenth-century explorers such as Zebulon Pike — receive less than 50 centimetres of rainfall a year. With boundless land but limited water resources, the region’s climate varies from desert to semi-arid.

## POWELL’S PLAN

In his 1878 plan, Powell called for the coordinated development of water, land, forestry and mineral resources. He also proposed legislative and social initiatives for implementing it. Rather than selling off rectangular plots of public land, as was usual at the time, he argued that lands should first be classified for uses such as irrigation, pasture, forestry or mining before being released for sale to settlers. Natural resources should be managed cooperatively at local and river-basin levels.

Maps and analyses of topography, geology and climate would be needed for the land



## OUTLOOK FOR EARTH

A Nature special issue on the IPCC  
[nature.com/ipcc2013](http://nature.com/ipcc2013)



classifications. Determining how much land to set aside for irrigation, for example, would require hydrologic calculations and engineering designs, as well as surveys of potential reservoir sites and supporting infrastructure plans. Financial feasibility would also need to be demonstrated for different locations.

Some of Powell's proposals were innovative — in particular his suggestion to tie together land and water rights. By grouping together land parcels, every farm could own its irrigated land, and communities could manage large tracts of unfenced pastureland. He believed that social institutions, such as schools and churches, and infrastructure, including roads and bridges, would spring up as a result.

Powell saw watersheds as the natural boundaries for resource management. Defining areas in which people's interests were interlinked and that were large enough to pay for irrigation projects, he believed that such great "watershed commonwealths" would oversee the development of resources within a drainage basin<sup>5</sup>.

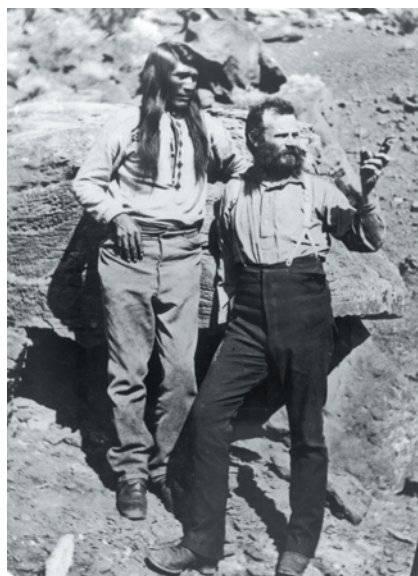
In 1881, Powell was appointed director of the US Geological Survey (USGS), tasked with surveying and classifying the lands of the west. He initiated a range of mapping and monitoring programmes, including streamflow, rainfall and preliminary reservoir surveys, culminating in a complete topographic map of the United States.

Mild weather conditions beginning after the conclusion of the American Civil War in 1865, and continuing through the early 1880s, allowed steady migration to the arid lands. States such as Nebraska, Kansas and Colorado doubled or tripled in population. But drought between the late 1880s and early 1890s, coinciding with poor national economic conditions, caused starvation, bankruptcy and social unrest<sup>6</sup>. Some 300,000 people migrated out of the arid region. The relative population drop of 50–75% was worse than for any subsequent drought<sup>6</sup>.

A national debate about the future of the region began, with Powell's plan at the centre. US Congress enacted legislation in 1888 that halted the release of public lands to settlers for two years while the USGS characterized the region. Powell estimated that it would take 10 years and US\$6 million to map the approximately 2.6 million square kilometres of arid lands; at the time, the USGS annual budget was just \$700,000. However, his first task was to recruit and train staff.

When Powell sought extra time to perform the full land and resource assessment, Congress and the public turned against him, and his plan fell from favour. The policy

**"Motivation to act might come only from rare catastrophic events such as droughts."**



Geologist John Wesley Powell (right) with his Paiute guide Tau-gu in Utah in 1873.

was reversed and the release of public lands recommenced in 1890. The USGS budget was frozen and its hydrographers discharged. With Powell's departure from the agency in 1894, attempts to develop large-scale resource management in the region effectively ended<sup>7</sup>.

But there was a continuing clamour for a national policy to address the climate of the arid lands, especially through irrigation. In 1896, Congress commissioned a report<sup>8</sup> to identify the best body to build and operate irrigation systems; basin-based organizations were thought to be too small to raise enough capital or to operate across states. The federal government was nominated, and the Reclamation Act of 1902 established a dominant federal role in reservoir construction.

#### POWELL'S LEGACY

The trajectory of Powell's plan holds lessons for the climate-change debate today. The parallels are striking — the arid-lands debate extended over a long period and resulted in scientific controversy and heated political discussions.

Powell established that data, analyses and proposals must be the basis for policy decisions. His assessment provided the first detailed understanding of the climate in the arid lands, as well as "definite information upon which the committees of Congress and the individuals interested could base their statements or conclusions"<sup>9</sup>. It left a broad scientific legacy, including tools such as a national topographic map and climate-monitoring systems that are still in use. Likewise, the IPCC reports will continue to provide value in decades to come through their monitoring, analyses and policy proposals.

The public response to Powell's plan and to the need to address the climate of the arid lands is telling. Opinion swung back and

forth over many decades as a multitude of factors intervened, including misinformation campaigns and external events such as economic recessions and unusual weather conditions. In a similar way, the climate-change debate today ebbs and flows.

The collapse of Powell's watershed-based strategy demonstrates the importance of choosing appropriate institutions and economic policies. His call for a class of self-financed regional-management entities, taking responsibility for resource management away from existing federal and state bodies, failed because it was not feasible economically or politically. Similarly, management scale and political realities must be considered when implementing various climate-change-mitigation strategies.

For complex issues such as climate change, history reminds us that the first comprehensive policy adopted will not be the last. Policies evolve. The addressing of irrigation needs through the Reclamation Act, although itself the culmination of a lengthy debate, was bolstered three decades later by soil-conservation efforts in response to the extreme Dust Bowl conditions of the 1930s.

Sadly, in the end, the motivation to act might come only from rare catastrophic events such as droughts; human cost offers a sharp impetus for politicians. As a US senator noted in a 1935 debate, a Great Plains dust storm that travelled more than 1,600 kilometres to reach the country's capital was "the most tragic, the most impressive lobbyist" in those earlier deliberations over climatic disaster<sup>10</sup>. ■

**K. John Holmes** is associate director of the Board on Energy and Environmental Systems at the US National Research Council in Washington DC.  
e-mail: jholmes@nas.edu

1. Gilpin, W. *The Central Gold Region* (Sower, Barnes and Company, 1860).
2. Powell, J. W. *Report on the Lands of the Arid Region of the United States* (US Department of the Interior, 1878).
3. USGS. *Twelfth Annual Report of the United States Geological Survey to the Secretary of the Interior, 1890–91, Part II—Irrigation* (Government Printing Office, 1891).
4. Holmes, K. J. & Wolman, M. W. *Environ. Mgmt* **27**, 177–193 (2001).
5. Powell, J. W. *The Century Magazine* **40**, 111–116 (1890).
6. Warrick, R. A. *Climatic Constraints and Human Activities* (eds Ausubel, J. & Biswas, A. K.) (Pergamon, 1980).
7. Van Hise, C. R. *The Conservation of Natural Resources in the United States* (Macmillan, 1910).
8. United States Army Corps of Engineers. *Preliminary Examination of Reservoir Sites in Wyoming and Colorado* (US Congress, 1897).
9. USGS. *Twenty-third Annual Report of the Director of the United States Geological Survey to the Secretary of the Interior 1901–2* (Government Printing Office, 1902).
10. Statement of Senator Thomas Gore (Oklahoma). *Congressional Record* **79**, 6013 (Government Printing Office, 1935).



Tarnow is walking her kids to school in Manhattan on the first day of spring, stopping to admire the sunlight shining through a pear tree's blossom. The lapsed-licence fraud investigator is about to be drawn into a sinister web of intrigue. An old acquaintance asks her to investigate the suspiciously successful dot-com for which he is filming a corporate documentary. Poking around brings Tarnow into contact with shady hackers, gregarious Italian-American venture capitalists, Mossad agents, bloggers, petty fraudsters who are in over their heads, trophy wives, a private investigator whose primary tool is his sense of smell, a pair of disarmingly likeable Russian gangsters with a fondness for hip hop, and a mysterious government operative. Some will be exiled, some will run away and some will carry on; not all will survive.

Pynchon often turns to science and engineering as sources of imagery and symbolism. His first novel, *V.*, which has its 50th anniversary this year, discussed circuit diagrams and humanoid automatons. *Gravity's Rainbow*, published 40 years ago, featured thermodynamics, chemistry and probability theory. Set in the eighteenth century, the 1997 *Mason & Dixon* meditated on the metaphysics of a clockwork universe while anachronistically nodding to the collapse of the wave function in quantum mechanics as a metaphor for the march of the British surveyors Charles Mason and Jeremiah Dixon westward through the American colonies into unknown territory. *Bleeding Edge*'s big-data-era backdrop is a perfect fit for Pynchon's sensibility.

When you are interested in paranoia and conspiracies, it is important to distinguish signals from noise, and true clues from false leads. This is also the situation in which modern scientists find themselves, searching for significance among countless spurious correlations in a flood of data.

Tarnow, when not breaking into secret passages or pulling her handgun on rapacious dot-com capitalists, spots fraudulent bookkeeping through deviations from Benford's Law (the expected distribution of first digits in numerical data), and her secretary's

observations of a pattern in the empty cells of an Excel spreadsheet.

Geek culture is evocatively rendered, often through throw-away jokes and asides. The interior of a smoke-filled limousine has "picked up a Gaussian blur"; a couple, asked if they will be living together

## FICTION

# Silicon and surveillance

Sean Carroll finds Thomas Pynchon on compelling form in a tale of big data and bigger conspiracies.

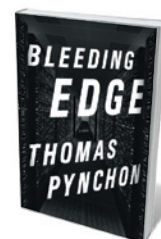
Thomas Pynchon's novels have several recurring themes: paranoia and conspiracy, pastiches of high and low culture, synchronicity and coincidence, shadowy networks lurking around every corner, and the impact of science and technology. With the coming of the Internet age and the surveillance society that sprang up in the wake of 11 September 2001, it seems as though reality has finally caught up with his vision. In his latest work, *Bleeding Edge*, Pynchon takes full advantage of this convergence.

The first question asked of a new Pynchon book is: is this one of the sprawling, spiralling, time-tripping monsters with innumerable characters and a plot that is tricky to bring into focus, like *Gravity's*

*Rainbow* or *Against the Day*; or is it one of the fun detective stories with a well-defined protagonist, like *The Crying of Lot 49* or *Inherent Vice*? *Bleeding Edge* is definitely in the latter category. There is a colourful cast of memorable personalities, and high jinks often ensue, but the tale is told linearly, from the point of view of an acknowledged main character, with something approximating an explicit goal.

The year is 2001. The dot-com bubble has just burst and Silicon Alley, New York's version of Silicon Valley, is in disarray. The Internet revolution is just beginning to gather steam. And, of course, the imminent 11 September attacks loom over every page.

The novel begins simply, in the mundane beauty of an ordinary morning. Maxine



**Bleeding Edge**  
THOMAS PYNCHON  
The Penguin Press:  
2013.



or separately, answer “True” in unison, and then explain “Sorry. Nonexclusive ‘or.’” Characters explore the Deep Web — a real but murky sector of the Internet unindexed by search engines and populated by dynamic content and entrepreneurs concerned with privacy. And they lurk in DeepArcher (say it out loud), a virtual space of Pynchon’s invention. This is a virtual-reality interface that grows organically, allowing its denizens to construct sanctuaries away from the surface web with no clear instructions — “part of getting constructively lost”. Characters who have met untimely ends in “meatspace” seem to linger on in DeepArcher, although it is impossible to tell who is really behind the face of an avatar.

The attacks on the World Trade Center happen about two-thirds of the way through the book. Rather than inducing a dramatic change, they lead to a slight shift in balance, pulling some notes down and bringing others to the foreground. Every type of conspiracy theory is tossed into the pot and stirred. Pynchon highlights the life-goes-on aspect of New York after the attacks, but laments the swiftness with which “forces in whose interests it compellingly lies to seize control of the narrative” work to keep people “cranked up, scared, and helpless”.

*Bleeding Edge* is an elegiac yet compulsively readable novel. The humour crackles, eliciting chuckles on almost every page. No one works magic with words like Pynchon, and here he is at the height of his powers, by turns gripping, thought-provoking, inventive, touching and poetic, not to mention warmly human. Tarnow is a rich, believable character, and we are fortunate to be privy to her wry commentary on the rogue’s gallery of characters who cross her path.

‘Big data’ is a modern buzzword, but a long-standing theme for Pynchon has been how the search for signals in the noise is, in part, a quest for meaning amid chaos and entropy. At the end of the novel, in a world transformed by atrocities large and small, Tarnow’s boys once again head off to school. This time they make the journey without their mother. Life does go on, even as everything changes; the best we can do is care. ■

**Sean M. Carroll** is a theoretical physicist at the California Institute of Technology in Pasadena. His books include *The Particle at the End of the Universe*.  
e-mail: seancarroll@gmail.com



A worker using an electrolysis furnace to produce aluminium in Krasnoyarsk, Russia.

#### INNOVATION

## Superpowered invention

Leonid Gokhberg and Dirk Meissner compare accounts on the trajectory of innovation in two towering economies.

Many of us believe that science is indispensable for generating innovation, and that innovation is a basis for manufacturing, which, in turn, is a must for the economy. But is the road that straight? In truth, it takes decades to recognize the fruits of scientific discovery and economic activity.

Science historian Loren Graham and environmental scientist Vaclav Smil examine this issue in two nations with economic and political models that have evolved very differently. Graham’s *Lonely Ideas* tackles Russia’s powerful history of scientific invention, its long-standing inability to benefit from this, and its post-Soviet potential for change. Smil’s *Made in the USA* looks at the United States’ innovation-led economic power, built on twentieth-century might in manufacturing everything from cars to electronics, but now weakening — and even creating trade deficits in high-tech products such as computer equipment.

What emerges are distinct challenges. The United States, Smil argues, should revitalize manufacturing (of household appliances or construction equipment, for example) to support growth and investment in health care and education, among other positive socio-economic impacts. In Russia, Graham shows, the challenge is to reshape an economy now

#### *Lonely Ideas: Can Russia Compete?*

LOREN GRAHAM  
MIT Press: 2013.

#### *Made in the USA: The Rise and Retreat of American Manufacturing*

VACLAV SMIL  
MIT Press: 2013.

reliant on oil and gas, and to make it competitive through technological upgrading and integration into global value chains — the interlinked processes that take a product from conception to end use.

Smil notes that US innovation co-evolved with the emergence of a middle class rooted in industry, such as steel manufacture in the nineteenth century, and automobile production in the twentieth. A Russian middle class has never fully emerged, yet Graham fails to tackle this issue. Rather, he focuses on the inventiveness of Russian scientists, describing impressionistic examples such as the early work of Nikolai Basov and Alexander Prokhorov on lasers in 1954. In exploring the country’s failure to turn invention into innovation, he briefly covers key factors such as the investment climate, institutional frameworks, policies and the societal mindset. His discussion of various areas of technology such as genetics is interesting, although some of his conclusions are odd. For instance, ►

► he blames the crash of a Sukhoi Superjet 100 during a 2012 demonstration flight on systemic failures rather than pilot error, subsequently shown to be the cause.

Both Graham and Smil analyse human resources and education as sources of growth. Smil reveals that the traditional distinction between blue- and white-collar jobs is diminishing in the United States, pushing up the level of qualifications needed. And he shows that although US universities attract the best talent globally, the country's overall education system is failing to train enough qualified individuals. Some universities have thus been forced to give remedial courses in subjects such as mathematics to first-year students.

In Russia, Graham suggests, attempts to regenerate the research sector by attracting high-level scientists, upgrading equipment and making greater use of talented students are providing the basis for innovation. The first signs of high-tech entrepreneurship are appearing. A 2010 government directive supporting university innovation has spawned multiple success stories, such as spin-off companies and growth in private-venture investment. Yet new initiatives, Graham argues, do not overcome the barriers between science and innovation still inherent

**"No nation can survive solely on digital industry."**

in Russian society. He points to the number of Russian scientists who remain psychologically trapped in the Soviet tradition of keeping research separate from both enterprise and universities. Nevertheless, the next generation is becoming aware that application and commercialization can complement fundamental scientific research and education.

Meanwhile, in the United States, willingness to take risks and convert inventions into commercial propositions is falling off with the rising standard of living, although it was once the norm, Smil shows. Population mobility also enabled US innovation, he argues. Graham reveals this as another dissimilarity to the situation in Russia, where even highly educated citizens tend to "stay near where they were born". But this is slowly changing: 37% of Moscow-based university students now hail from elsewhere in Russia or abroad.

Approaches to growth, whether US or Russian, can backfire. Smil points to the drive of US businesses to maximize profits at short notice and arrives at a provocative question: will the United States be able to maintain its role as one of (if not the) leading economic powerhouses by relying on digital-age business models? That question remains open.

Taking *Lonely Ideas* and *Made in the USA* together, an overarching message emerges: innovation needs to be allied to tangible outcomes, to something that people can use. No nation can survive solely on digital industry;

a good living standard comes from combining innovation in the real and digital economies, and in services. But these demand clever policies and frameworks, including a favourable climate for competition and investment, property protection and the rule of law.

Both the United States and Russia are beginning to actively tackle their national innovation challenges. The leap in the US chemical industry based on shale gas and the modernization of the fuel sector and automotive industry in Russia seem to hold promise — but things may look very different in the coming decade. Sustainable growth in both countries will probably come from highly automated manufacturing, which will demand a better educated and trained labour force. The digitization of manufacturing, encompassing robots, three-dimensional printing and more, will continue apace along with the need for new skills. We will need, in short, to be ready for new twists in the road from lab to reality. ■

**Leonid Gokhberg** is first vice-rector of the National Research University - Higher School of Economics (HSE) in Moscow and editor-in-chief of the journal *Foresight-Russia*. **Dirk Meissner** is deputy head of the Laboratory for Science and Technology Studies at HSE Institute for Statistical Studies and Economics of Knowledge in Moscow. e-mail: lgokhberg@hse.ru

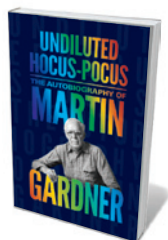
## MATHEMATICS

# Master puzzler

David Singmaster delights in the autobiography of Martin Gardner, whose *Scientific American* maths column enchanted tens of thousands.

For half a century, Martin Gardner (1914–2010) was an international scientific treasure. As the author of *Scientific American*'s Mathematical Games column for 25 years, he introduced many thousands to the pleasures of mathematics. He enchanted tens of thousands more with more than 100 books spanning everything from pseudoscience and magic to *Alice's Adventures in Wonderland*. To anyone who knows Gardner's work, his self-proclaimed "rambling autobiography" — the posthumously published *Undiluted Hocus-Pocus* — comes as a delightful surprise.

Gardner reveals the roots of his unusual mix of expertise in his childhood in Tulsa, Oklahoma. His father — a freelance oil prospector with a background in geology — taught Gardner basic science such as why



**Undiluted Hocus-Pocus: The Autobiography of Martin Gardner**  
MARTIN GARDNER (WITH PERSI DIACONIS AND JAMES RANDI)  
Princeton University Press: 2013.

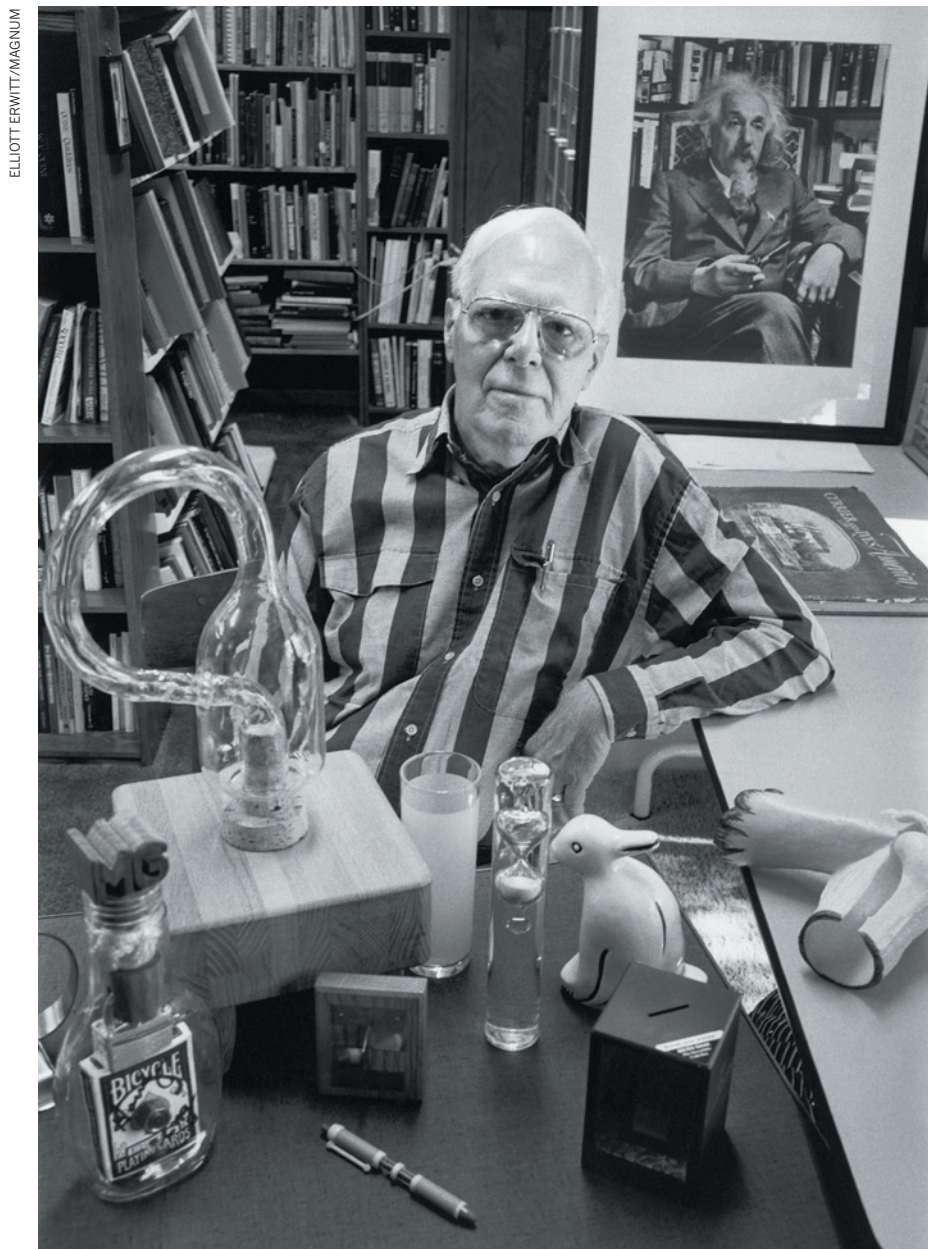
the Moon has phases, provided him with a small laboratory and taught him some magic tricks. Gardner learned to read by looking over his mother's shoulder as she read aloud L. Frank Baum's children's classic *The Wonderful Wizard of Oz* (1900). He subscribed to *Science and Invention* magazine and *Amazing Stories*, the first science-fiction magazine, launched in 1926. He performed his first magic trick at the age of

eight, later following the famous US Tarbell Course in Magic.

Gardner hated high school, except for mathematics and physics, noting that the "important history ... was the history of science". Here, he writes, he penned "lots of mediocre poetry" and invented 'cherchez la femme', a flexagon-type puzzle — flat paper models folded different ways to reveal various images. In 1934, when Gardner was just 20, *Hobbies* magazine published his article on collecting mechanical puzzles — the first of its kind.

He had wanted to study physics, but instead read philosophy at the University of Chicago in Illinois. Its new president, Robert Hutchins, fomented an educational revolution by appointing Mortimer Adler to a chair in philosophy without consulting





Martin Gardner, pictured in 1995.

members of the department, most of whom resigned. Hutchins and Adler went on to promote the Great Books scheme — a curriculum focused on texts by scientific and literary luminaries from Archimedes to Virginia Woolf — and a highly flexible undergraduate programme. Gardner enjoyed the ferment as philosophers came and went, and recalls seeing Enrico Fermi cycling to the university's Stag Field, where the great physicist was making the first atomic pile (an early reactor) in an underground squash court.

Gardner's complex of interests began to bear fruit in the 1950s, as he cut his writing teeth on journalism. He moved to New York to edit the quality children's monthly *Humpty Dumpty* for eight years. He published articles on maths and magic for

*Scripta Mathematica*; these were gathered together for his first recreational mathematics book, *Mathematics, Magic and Mystery*, in 1956.

That same year, Gardner was shown a hexahexaflexagon, a puzzle made by folding a length of paper into a hexagonal Möbius strip; it was the work of four students at Princeton University (one of whom was Richard Feynman). Gardner thought *Scientific American* might like an article on it and the piece kick-started his much-loved column. In it, Gardner introduced or popularized a vast range of ideas, including polyominoes (the shapes formed by joining squares edge-to-edge); the Soma cube; the superellipse; M. C. Escher's iconic images, such as the *Endless Staircase*; Roger Penrose's tilings; and trapdoor ciphers

and public-key cryptography, the basis of all financial transactions on the Internet. These pieces were eventually collected into 15 books.

Magic gripped Gardner throughout his life, and he may have written more on it than on mathematics. I feel that his years spent writing for children may partly account for the exceptional clarity and directness of his writing for adults. He became a leading expert on Lewis Carroll, with his *Annotated Alice* (Bramhall House, 1960) selling more than a million copies and explaining the mathematical, logical and literary associations in the book.

He also became a specialist on Baum, the prolific British writer G. K. Chesterton, and some minor poets. Gardner produced so much, on so many subjects, that it was rumoured that his name was the pseudonym of a writers' collective, such as the French mathematicians who published as Nicolas Bourbaki.

### DEBUNKING PSEUDOSCIENCE

Gardner received an immense amount of correspondence and apparently replied to all of it. Many students, encouraged by his friendly and prompt replies, became mathematicians.

Among scientists, Gardner is best known for his writing debunking pseudoscience. This began with 'The Hermit Scientist' in *The Antioch Review* of winter 1950–1951. In this, he described ideas such as psychiatrist Immanuel Velikovsky's catastrophist theories about ancient history as examples of pseudosciences created by persons working alone.

He expanded these ideas in books such as *Fads and Fallacies in the Name of Science* (Dover Publications, 1957) and *Science: Good, Bad and Bogus* (Prometheus, 1981). In 1976, Gardner helped to found the Committee for Scientific Investigation of Claims of the Paranormal (now the Committee for Skeptical Inquiry), which believes that it is "the duty of scientists to debunk bad science". Amen.

Gardner's passion for writing and his warmth and humour shine forth on every page of this book, making it a memoir of a great human being — a 'rational man', as Isaac Asimov had it. Almost all of Gardner's books are still in print. They stand as a remarkable testament to the independent scientific life. ■

**David Singmaster** is a retired professor of mathematics at London South Bank University and an honorary research fellow at University College London. He was the first person to write about the Rubik's Cube outside Hungary and devised the now standard notation for it.  
e-mail: zingmast@gmail.com

# Correspondence

## Yasuní oil plans call for vigilance

Ecuador's President Rafael Correa announced last month that development would begin in the eastern edge of the world's most biodiverse area, the Yasuní Biosphere Reserve. Despite his innovative 2007 proposal to leave oil deposits there untouched, insufficient offsetting funds materialized. Poverty applies the greatest pressure to nature, Correa declared.

Correa promised that the latest technologies will be used to minimize direct impacts and to avoid indirect ones altogether. He pointed out that sacrificing the sector, known as the ITT block, will affect "less than 0.1%" of the reserve's area and generate a US\$18.2-billion pay-off.

The devastating history of oil extraction in other parts of Ecuador might prompt scepticism about this figure, but it could be realized if offshore methods are faithfully incorporated and no new access roads are built. Everything possible must be done to preserve this unique piece of Amazonia, which contains more than 100,000 species in each hectare of intact forest.

However, contracts between the government and an operator for the development of an oil-concession area can be undermined by rapid advances in technology, so that even the most visionary operations end up using obsolete methods and equipment. Sensitive and biodiverse areas such as Yasuní need to retain new technologies throughout any extraction operation.

**Kelly Swing** *Tiputini Biodiversity Station, University of San Francisco de Quito, Ecuador.*  
kswing@usfq.edu.ec

## Ban resistant strains from food chain

No reliable treatment is available for humans infected with carbapenem-resistant

Enterobacteriaceae (CREs; see *Nature* **499**, 394–396; 2013). Because these antibiotic-resistant bacterial pathogens are already entering the food chain (J. Fisher *et al.* *J. Antimicrob. Chemother.* **68**, 478–480; 2013) and can be transmitted through oral consumption (A. R. Manges and J. R. Johnson *Clin. Infect. Dis.* **55**, 712–719; 2012), we call for a zero-tolerance ruling on CREs in retail food to stop the situation from getting out of control.

It was estimated that in 2007 more than 1,500 people in Europe died from an invasive infection caused by a strain of *Escherichia coli* that was resistant to third-generation cephalosporins and originated in poultry (P. Collignon *et al. Emerg. Infect. Dis.* **19**, 1339–1340; 2013). Mortality from CREs is expected to be much higher if it spreads in poultry on a comparable scale.

Current regulations on antimicrobial usage are largely ineffective — mainly because they operate only at a national level, whereas antimicrobial sales and consumption are global. An international ban on the sale of food items containing CREs could force imported and locally produced food to meet the same standards.

Such a ban would also comply with the guidelines of the Codex Alimentarius Commission, which sets international food standards under the World Trade Organization's Agreement on the Application of Sanitary and Phytosanitary Measures.

Given that CREs can be detected rapidly and reliably in culture-based or molecular assays, routine testing of meat products should be achievable. We urge politicians and decision-makers to protect public health by acting promptly, before the problem becomes widespread and unmanageable.

**Jan Kluytmans\*** *VU University Medical Center, Amsterdam, the Netherlands.*  
jankluytmans@gmail.com

*\*On behalf of 8 co-signatories (see go.nature.com/qahy4b for full list).*

## E-learning booster in developing world

Online courses and virtual education laboratories are proving valuable in the resource-limited developing world, where there are many more potential learners than can be accommodated by the pool of experienced classroom teachers and institutions offering scientific education (see *Nature* **495**, 160–163; 2013 and *Nature* **499**, 268–270; 2013).

India has developed its own innovative examples of freely available e-learning resources. Scientific-education materials, such as those supplied by the Sakshat Virtual Labs and the National Programme on Technology Enhanced Learning, are also available to a global audience (S. Ray *et al. PLoS Biol.* **10**, e1001353; 2012).

In developing countries, the anticipated benefits of such resources extend beyond distance learning. Science and technology skills can fuel a knowledge-driven economy, narrowing a country's poverty gap and leading to national prosperity that is embedded in sustainable practices.

**Sanjeeva Srivastava\*** *Indian Institute of Technology Bombay, Mumbai, India.*

sanjeeva@iitb.ac.in

*\*On behalf of 7 co-signatories (see go.nature.com/9zknx1 for full list).*

## Boycott challenges research tactics

Italy's National Agency for the Evaluation of the University System and Research (ANVUR) has released its evaluation report for the period 2004–10 (see *Nature* doi.org/nrx). The country's largest publicly funded research organization, the National Research Council (CNR), came unexpectedly

low in this ranking — despite its impressive record in basic and applied research products (see go.nature.com/kv2ipo; in Italian). In our view, this is a result of a boycott of the evaluation process by a group of CNR researchers.

During ANVUR's evaluation, a 'reform' under Silvio Berlusconi's government centralized decisions within the CNR by a board of five government nominees, excluding CNR researcher representatives — making the CNR the only public research organization in Italy without any researchers on its governing board.

In protest, some 15% of CNR researchers decided not to submit their publications for evaluation by ANVUR. Because missing products score negatively, the effect of the boycott was amplified and seems to have affected the CNR's ranking.

In a press release, the CNR's chairman, Luigi Nicolais, attributed the council's low ranking to bias in the ANVUR criteria, indicating that the CNR should help Italy's growth and support its enterprises and public administrations, and that these activities should in future be evaluated alongside scientific publications. Most CNR scientists would object to this, however, as they see research as their principal task.

Not all CNR researchers (including ourselves) agreed about the strategic value of the boycott, but it was a last stand for internal democracy, representation and transparency in decision-making at the CNR. To our knowledge, these issues remain unresolved.

**Umberto Amato** *CNR Institute for Applied Mathematics "Mauro Picone", Naples, Italy.*

**Patrizia Lavia** *Institute of Molecular Biology and Pathology, Rome, Italy.*

patrizia.lavia@uniroma1.it

**Vito Mocella** *Institute for Microelectronics and Microsystems, Naples, Italy.*



## The cause of the pause

A global climate model that factors in the observed temperature of the surface ocean in the eastern equatorial Pacific offers an explanation for the recent hiatus in global warming. [SEE LETTER P.403](#)

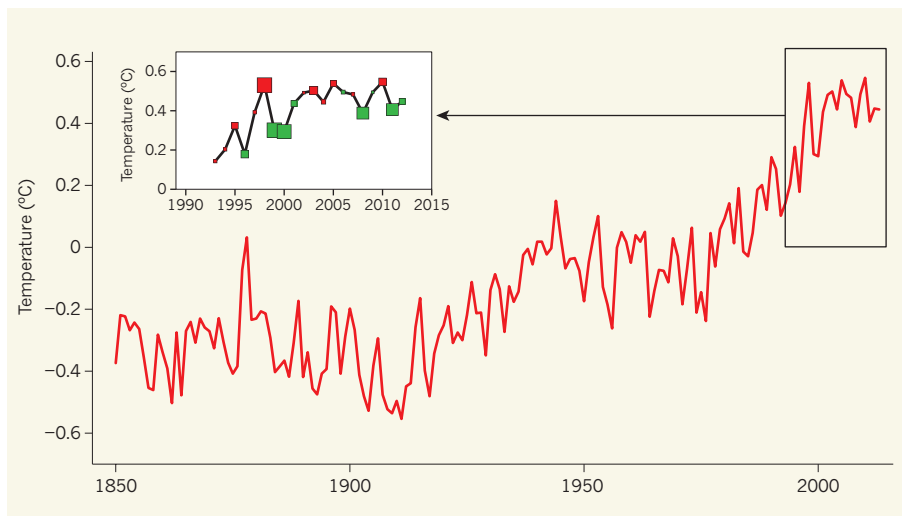
ISAAC M. HELD

After a rise of  $0.5^{\circ}\text{C}$  in the 25 years starting in the mid-1970s, the change in Earth's global mean surface temperature has been close to zero since the turn of the century (Fig. 1). This hiatus in global warming has occurred despite retreating Arctic sea ice and rising sea levels. On page 403 of this issue, Kosaka and Xie<sup>1</sup> make a strong case for the hypothesis that this hiatus is driven by the equatorial Pacific.

Although the rise in carbon dioxide and other greenhouse gases explains many aspects of the overall warming trend over the past century (including the heat uptake by the oceans and the spatial and seasonal patterns of the warming), it cannot explain the multi-decadal fluctuations superimposed on this trend (Fig. 1). Forcing agents such as anthropogenic and volcanic aerosols and variations in the Sun's energy output are often called on to explain these features. Internal fluctuations of the atmosphere–ocean system, the low-frequency tail of our chaotic weather, could also be responsible. It is likely that both forced and internal variations have a role.

A mix of explanations has been offered for the recent hiatus: the minimum in solar energy output in the latest 11-year sunspot cycle lasted longer than usual<sup>2</sup>; stratospheric water vapour, which warms the surface, has been relatively low since 2000<sup>3</sup>; and the El Niño–Southern Oscillation (ENSO) cycle of warm El Niño and cold La Niña phases in the equatorial Pacific, which is known to affect global mean temperature and not just the temperature of the equatorial Pacific, has favoured the La Niña phase since the major El Niño event of 1997–98. Empirical models that fit the observed hiatus have generally relied on La Niña-related cooling to offset a large fraction of the greenhouse-induced warming<sup>4,5</sup>. Consistently, analyses of the heat being taken up by the oceans have pointed to an increase in this heat uptake, predominantly in the Pacific, as underlying the hiatus<sup>6,7</sup>.

Kosaka and Xie base their case on simulations using a global climate model<sup>8</sup>. A single run of the model, starting with initial conditions at a time well before the period of interest, generates a sample of the model's internal



**Figure 1 | Temperature evolution.** The graph shows the global mean surface temperature relative to the 1961–90 mean, based on the HadCRUT4.2.0.0 data set<sup>16</sup>. The inset shows the 1993–2012 time span, with green denoting La Niña years and red, El Niño years; the size of the symbol indicates the strength of La Niña/El Niño according to the Ocean Niño Index<sup>17</sup>. (The Niño index for year  $N$  is computed by averaging from October of year  $N-1$  to September of year  $N$  to account for the lag between the El Niño–Southern Oscillation and global mean temperature.) Kosaka and Xie<sup>1</sup> argue that a cooling trend in a region covering only about 8% of Earth's surface in the eastern equatorial Pacific, a trend associated with a preponderance of La Niña events, explains the absence of global mean warming over the most recent decade.

variability, its chaotic weather on all timescales, superimposed on its response to the imposed forcing agents. By averaging over many runs of the model, each starting with different initial conditions, the forced climate change can be separated from the model's internal variability. In its forced response, the model warms globally by about  $0.2^{\circ}\text{C}$  in the most recent decade. Kosaka and Xie then intervene in the model to force the temperature of the surface ocean in the eastern equatorial Pacific to follow observations. The influence of this constraint propagates through the system, compensates for the  $0.2^{\circ}\text{C}$  warming and results in an excellent fit to the observed evolution of global mean temperature.

The flat annual-mean-temperature trend during the hiatus consists of distinct cooling centred in the Northern Hemisphere winter, especially over land<sup>9</sup>, and warming or little change in other seasons. This seasonal cycle of global temperature trends is roughly captured by the constrained model, providing further support for the central influence

of the equatorial Pacific in the hiatus.

A trend towards drier conditions in the southern continental United States is a feature of most climate-model responses to increasing greenhouse gases. One might guess that global mean cooling forced by a La Niña-like trend that compensates for greenhouse warming for 10–15 years might also cancel out this drying tendency. But the opposite is the case: La Niña is known to favour drying in this same region by displacing the Pacific storm track polewards. Generating a hiatus in global warming through a La Niña-like trend should exacerbate rather than ameliorate drying due to greenhouse gases in this region, as is seen in the constrained model of Kosaka and Xie. The hiatus in global warming and the prolonged US drought over this same time period are probably closely related.

A central question for the future, not directly addressed by Kosaka and Xie, is whether the recent trend towards a more La Niña-like state is itself forced or a result of internal variability. Models have provided little support for

viewing this trend in the equatorial Pacific as being forced, although there are suggestions for how such a response might be possible<sup>10,11</sup>. The alternative hypothesis, that the equatorial cooling trend is due to internal variability of the climate system, is gaining support from modelling efforts<sup>12</sup>.

Expectations for global temperature change in the next decade depend on whether recent trends in the equatorial Pacific are viewed as being due to a coincidental combination of chaotic El Niño and La Niña events, or as having connections to a mode of variability with deeper roots and intrinsically longer timescales. This deeper variability could, for example, affect the subsurface pathways taken by waters upwelling in the eastern equatorial Pacific (as may have occurred in the mid-1970s shift towards a state more favourable to El Niño events<sup>13</sup>). A pattern of variability in the North Pacific known as the Pacific Decadal Oscillation<sup>14</sup> (PDO) is one candidate for introducing these lower-frequency internal variations into the equatorial Pacific. But the ability of the PDO to drive variability in the tropics, rather than simply being driven by the tropics, remains unclear<sup>15</sup>.

The uncertainty in climate sensitivity, a measure of the magnitude of the forced temperature response, is a major source of the spread in climate projections over the twenty-first century. Attempts at constraining climate sensitivity with the observed warming are affected by how internal variability is separated from the forced response in these observations. Manipulations of climate models similar to that described by Kosaka and Xie may allow this separation to be performed more convincingly and thereby reduce the uncertainty in climate sensitivity. Characterizing the low-frequency tail of internal climate variability is of paramount importance, whether we are interested in the next 10 or the next 100 years. ■

**Isaac M. Held** is in the Geophysical Fluid Dynamics Laboratory, National Oceanic and Atmospheric Administration, and the Program in Atmospheric and Oceanic Sciences, Princeton University, New Jersey 08542, USA.  
e-mail: isaac.held@noaa.gov

- Kosaka, Y. & Xie, S.-P. *Nature* **501**, 403–407 (2013).
- Frölich, C. *Surv. Geophys.* **33**, 453–473 (2012).
- Solomon, S. *et al.* *Science* **327**, 1219–1223 (2010).
- Lean, J. L. & Rind, D. H. *Geophys. Res. Lett.* **36**, L15708 (2009).
- Foster, G. & Rahmstorf, S. *Environ. Res. Lett.* **6**, 044022 (2011).
- Watanabe, M. *et al.* *Geophys. Res. Lett.* **40**, 3175–3179 (2013).
- Balmaseda, M. A., Trenberth, K. E. & Källén, E. *Geophys. Res. Lett.* **40**, 1754–1759 (2013).
- Delworth, T. L. *et al.* *J. Clim.* **19**, 643–674 (2006).
- Cohen, J. L., Furtado, J. C., Barlow, M., Alexeev, V. A. & Cherry, J. E. *Geophys. Res. Lett.* **39**, L04705 (2012).
- Clement, A. C., Seager, R., Cane, M. A. & Zebiak, S. E. *J. Clim.* **9**, 2190–2196 (1996).

- Liu, J., Wang, B., Cane, M. A., Yim, S.-Y. & Lee, J.-Y. *Nature* **493**, 656–659 (2013).
- Meehl, G. A., Arblaster, J. M., Fasullo, J. T., Hu, A. & Trenberth, K. E. *Nature Clim. Change* **1**, 360–364 (2011).
- Guilderson, T. P. & Schrag, D. P. *Science* **281**, 240–243 (1998).

- Mantua, N. J. & Hare, S. R. *J. Oceanogr.* **58**, 35–44 (2002).
- Newman, M., Compo, G. P. & Alexander, M. A. *J. Clim.* **16**, 3853–3857 (2003).
- www.metoffice.gov.uk/hadobs/hadcrut4
- www.cpc.noaa.gov/products/analysis\_monitoring/ensostuff/ensoyears.shtml

## DEVELOPMENTAL NEUROSCIENCE

# Miniature human brains

**A study shows that stem cells can be used to generate self-organizing three-dimensional tissues that mimic the developing human brain. These tissues provide a tool for modelling neurodevelopmental disorders. [SEE ARTICLE P.373](#)**

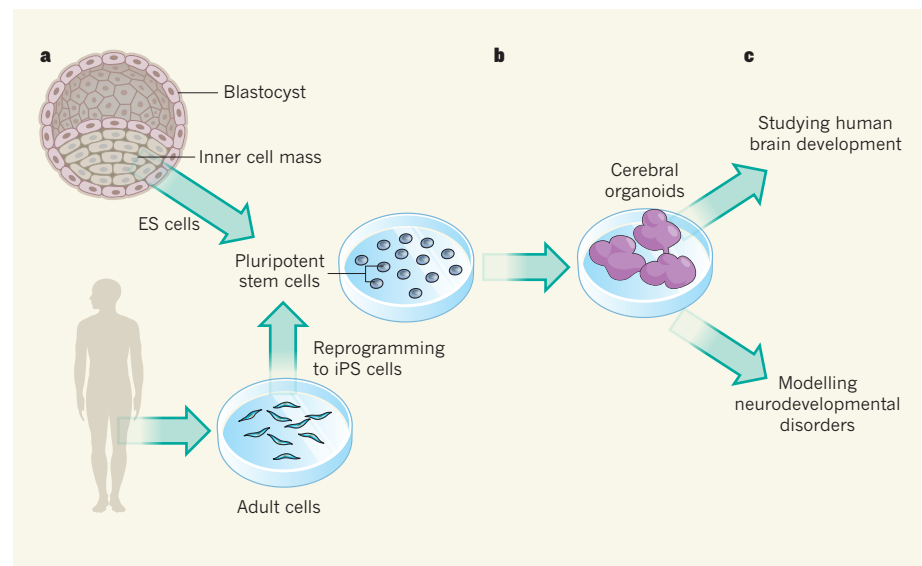
OLIVER BRÜSTLE

**P**luripotent stem cells can differentiate into almost all cell types of the body. This process can also be recapitulated *in vitro*, using both induced pluripotent stem (iPS) cells, which are themselves generated through the reprogramming of differentiated cells, and embryonic stem (ES) cells. But on page 373 of this issue, Lancaster *et al.*<sup>1</sup> demonstrate that pluripotent stem cells are more than only a source of different cell types. The authors find that, when left to aggregate under appropriate conditions, neural cells generated from human iPS cells self-organize into cerebral organoids

— up to pea-sized miniatures of developing human brain tissue. These structures are not just peculiar lab artefacts. As the authors show, the organoids recreate early steps in the formation of the human brain's cerebral cortex, and so lend themselves to studies of brain development and neurodevelopmental disorders\*.

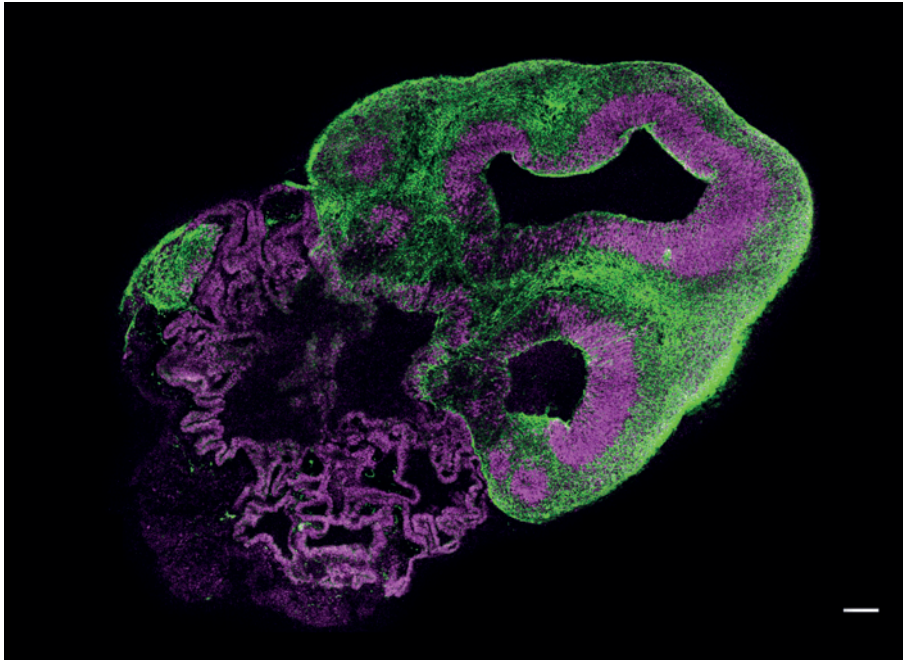
The idea that early neural cells generated from pluripotent stem cells can recapitulate neurodevelopmental programs has built up gradually over the past five years. In 2008, researchers reported<sup>2</sup> that mouse ES cells

\*This article and the paper under discussion<sup>1</sup> were published online on 28 August 2013.



**Figure 1 | Organoid generation from human pluripotent stem cells.** Pluripotent stem cells can be derived as either embryonic stem (ES) cells from the inner cell mass of the blastocyst (70–200-cell embryos) or induced pluripotent stem (iPS) cells through the reprogramming of adult cell types<sup>9</sup>. Lancaster *et al.*<sup>1</sup> show that neural cells derived from pluripotent stem cells and left to aggregate under appropriate culture conditions in suspension self-organize into cerebral organoids with tissue architectures that are reminiscent of the human cerebral cortex. Such organoids can serve as a reductionist three-dimensional model for studying mechanisms of both early human brain development and neurodevelopmental disorders.





**Figure 2 | A detailed view.** Image of an organoid generated by Lancaster *et al.*<sup>1</sup>, fluorescently labelled for neurons (green) and progenitor cells (magenta). The morphology of the upper-right region matches that of the cerebral cortex, and the region in the lower left has a similar morphology to that of the choroid plexus. (Scale bar, 100  $\mu\text{m}$ .)

that were allowed to differentiate on plastic dishes generated ‘waves’ of neurons. This process resembled the sequential birth of different classes of neuron that gradually form the developing cerebral cortex, layer by layer, during normal development. At the same time, another group provided stunning data<sup>3</sup> showing that neural cells derived from ES cells aggregate in suspension and can generate intriguing architectures of concentric cell layers that express molecular markers of different cortical layers. More recently, the same group demonstrated the formation of stretches of cortex-like neuroepithelial cell layers and of entire optic cups (primitive eye structures), as well as stratified retinas from mouse and human ES cells<sup>4–6</sup>.

Lancaster *et al.* take these fascinating observations to the next level (Fig. 1). Applying methods typically used in developmental neurobiology, they provide unprecedented detail on how closely these self-organizing aggregates match human-specific aspects of brain development. The cerebral organoids (as the authors call them) grow to up to 4 millimetres in size and can be maintained for several months in a spinning bioreactor, which improves nutrient absorption. Moreover, they contain a fluid-filled cavity (similar to the ventricles that carry cerebrospinal fluid within the human brain) and even structures resembling the choroid plexus — a cauliflower-like tissue found inside the ventricles, which in the brain generates the cerebrospinal fluid (Fig. 2). In the organoids, the tissue surrounding this ventricle segregates into distinct areas expressing markers of different brain regions.

Remarkably, the organoids also contain striking architectures reminiscent of the developing human cerebral cortex, which is evolutionarily the most complex tissue in the animal kingdom. As in the native brain, these cortical areas segregate into different layers. In the innermost layer — the equivalent of the ventricular zone — radial glial cells divide and give birth to neurons. The developing human brain also uses a second, afterburner-like matrix called the outer subventricular zone, where progenitors derived from radial glial cells in the ventricular zone continue to divide to generate the bulk of the neurons required to build up the large cerebral cortex. Intriguingly, this peculiarity of human cortical development is also well recapitulated in Lancaster and colleagues’ organoids.

On the basis of their own cell-aggregation studies, other researchers previously proposed<sup>7</sup> a multistep process of self-organization, whereby distinct cohorts of cells generated by self-assembly start to interact with each other to engage in self-patterning and self-driven morphogenesis. Lancaster and co-workers’ data are consistent with this, pointing to potential interactions between the spatially segregated cell populations within the aggregate. Distinct areas dispatch cohorts of neurons to travel to and populate the cortex-like structures — activities similar to those seen in the developing brain when interneurons generated in the ventral part of a brain region known as the telencephalon embark on a long route of tangential migration to populate the dorsal cortical areas.

Despite these compelling data, the

realization of a ‘brain in a dish’ remains out of reach. Within the organoids that Lancaster *et al.* describe, areas mimicking different brain regions are randomly distributed and lack the shape and overall spatial organization seen in the brain. Moreover, because they are devoid of a circulatory system, the metabolic supply of such organoids is restricted, which means that they can grow to only a few millimetres in size. Even then, their core represents a dead zone of cells starved of oxygen and nutrients. And although areas of these organoids’ exteriors clearly bear a resemblance to the developing cerebral cortex, it remains unclear whether they can advance to the complex six-tiered architecture of their natural counterpart.

Yet, even with these limitations, the potential applications of such an *in vitro* model are fascinating. For the right questions, cerebral organoids could provide a reductionist model with which to study mechanisms of early human brain development under controlled conditions in a laboratory dish. Such studies can even be extended to neurodevelopmental disorders. Recently, for example, a study<sup>8</sup> showed that layered cortex-like architectures in suspension culture can be generated from human iPS cells, opening up the possibility of establishing patient-specific organoids.

Lancaster *et al.* take that leap and generate cerebral organoids using iPS cells derived from skin fibroblasts of a patient with a genetic form of microcephaly — a neurodevelopmental disorder characterized by substantial reduction in brain size that is difficult to study in mouse models because of species-specific differences in brain development. When the authors compared the patient-derived organoids with those derived from healthy controls, they found fewer progenitor cells and an increased number of differentiated neurons. This observation points to premature neuronal differentiation as the mechanism underlying this form of microcephaly. The resolution of this system for analyses of more complex brain pathologies remains to be determined. The current paper, however, clearly puts neural aggregation cultures on the map of research tools for both developmental biology and biomedicine. ■

**Oliver Brüstle** is at the Life & Brain Centre, University of Bonn, 53105 Bonn, Germany. e-mail: brustle@uni-bonn.de

1. Lancaster, M. A. *et al.* *Nature* **501**, 373–379 (2013).
2. Gaspard, N. *et al.* *Nature* **455**, 351–357 (2008).
3. Eiraku, M. *et al.* *Cell Stem Cell* **3**, 519–532 (2008).
4. Eiraku, M. *et al.* *Nature* **472**, 51–56 (2011).
5. Nasu, M. *et al.* *PLoS ONE* **7**, e53024 (2012).
6. Nakano, T. *et al.* *Cell Stem Cell* **10**, 771–785 (2012).
7. Sasai, Y. *Nature* **493**, 318–326 (2013).
8. Mariani, J. *et al.* *Proc. Natl Acad. Sci. USA* **109**, 12770–12775 (2012).
9. Takahashi, K. *et al.* *Cell* **131**, 861–872 (2007).

## OCEANOGRAPHY

# Mountain waves in the deep ocean

**Internal lee waves are a player in ocean dynamics that may make an important contribution to deep-ocean mixing. They warrant serious consideration for inclusion in the next generation of climate models.**

JENNIFER MACKINNON

**T**urbulent mixing in the ocean interior plays a crucial part in driving downward transport of heat and dissolved greenhouse gases and upward transport of biologically essential nutrients. It also ultimately supplies the deep ocean with the energy that drives a global network of currents known as the meridional overturning circulation. Climate models that do not represent this mixing appropriately will be unable to reproduce present or future climate accurately<sup>1</sup>. Many previous studies have focused on the breaking internal waves that are driven by tides and winds as the dominant source of that turbulence. In a paper published in *Geophysical Research Letters*, Nikurashin and Ferrari<sup>2</sup> describe a previously under-appreciated dynamic mechanism — internal lee waves — that may significantly contribute to mixing in the deep ocean.

Away from the direct influence of surface forcing, most turbulent mixing in the ocean interior is driven by breaking internal gravity waves<sup>3</sup>. These propagate along and across density interfaces within the ocean, similar to the interfacial waves you might see between oil and vinegar in a glass or even between coffee and milk in a well-made cappuccino. Compared with the more familiar surface waves, internal waves are much slower, with periods of hours instead of seconds. Their breaking is in many ways analogous to that of surface waves on the beach, albeit in a slow-motion, larger-than-life way — the wave height may reach tens or even hundreds of metres<sup>3</sup>. But unlike waves at the beach (which lose all of their energy to heat or sound with each thunderous crash), some of the energy lost by breaking internal waves increases the potential energy of the ocean by mixing stratified water and raising its centre of mass. It is this potential energy that is eventually converted into the kinetic energy of the meridional overturning circulation.

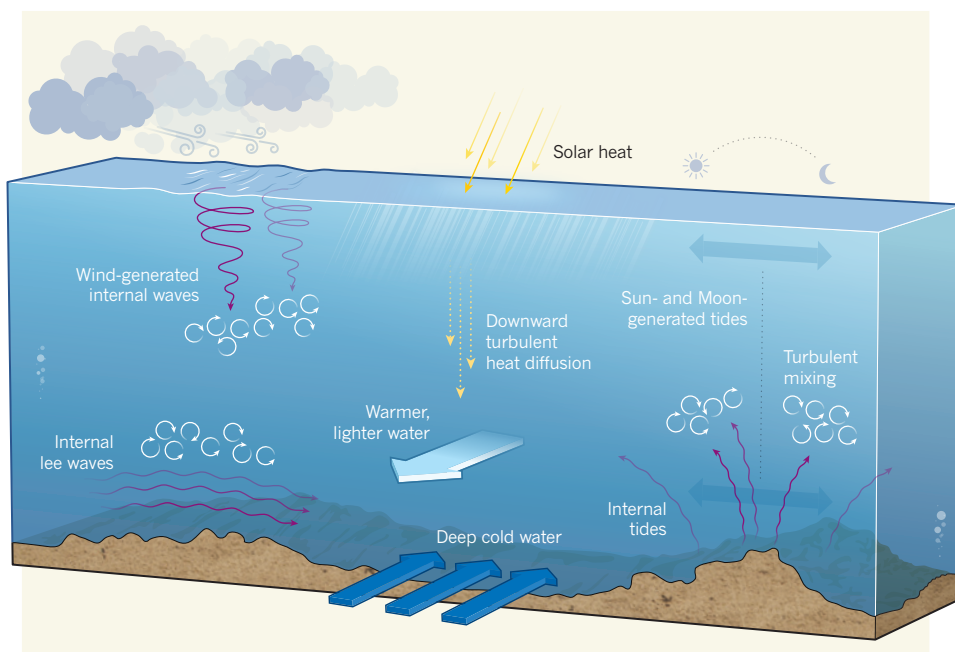
Over the past decade, most of the focus of oceanographers has been on the geography and life cycle of internal waves created by winds and tides<sup>4</sup>. In the deep ocean, internal waves with tidal frequencies have been assumed to be a dominant mechanism for turbulent mixing. These internal tides are

produced when the surface tide, generated by the Sun and Moon, forces dense water up and over sea-floor topography. This happens in the same way that tides pull and push water up and down the beach once or twice a day. As water goes back and forth, up and over, it perturbs the normally flat interfaces between density layers known as isopycnals, and creates internal waves along those surfaces. Some of that energy dissipates locally, producing a pattern of enhanced turbulence over rough topography (Fig. 1). Recent analysis<sup>5</sup> of climate models has shown that modelled ocean circulation and heat content are sensitive not only to the average level of turbulent mixing in the deep ocean, but also to its detailed geography.

In their study, Nikurashin and Ferrari discuss a related but relatively new player on the scene — internal lee waves. These are internal waves that are produced by comparatively steady flow over sea-floor topography. As

with internal tides, dense water is forced up and over topographic obstacles and thus perturbs density surfaces. The idea is similar to the standing waves that delight river kayakers, or to atmospheric mountain waves over abrupt geographical features — such as Mount Rainier in Washington — that are often visible as lenticular clouds. Although these waves seem stationary, energy propagates up into the overlying moving fluid. When the flow speed changes or waves become large, these internal lee waves can also break and produce turbulent mixing (Fig. 1). In the atmosphere, inclusion of the dynamic consequences of mountain waves is known to be important for accurate weather forecasts.

Nikurashin and Ferrari argue that internal lee waves may be prominent in places where strong deep flows encounter rough topography. In particular, their theoretical calculations highlight the Southern Ocean, where the strong Antarctic Circumpolar Current flows over the rough topography of Drake Passage, which lies between South America and the Antarctic Peninsula. They propose that turbulent mixing associated with breaking internal lee waves, when globally integrated, contributes one-third of the mixing required to transform the deepest dense water into less-dense water. Using formulations for mixing induced by lee waves and by internal tides, they calculate the total rate of water-mass transformation to be 20–30 sverdrups (1 sverdrup is  $10^6 \text{ m}^3 \text{ s}^{-1}$ ), which is of the same order as that calculated by other measures of the deepest component of



**Figure 1 | Ocean mixing.** Turbulent mixing (curly arrows) is driven by breaking internal waves in the ocean interior. It transports solar heat downwards from the surface to the abyss and transforms the deepest, coldest waters into warmer, less-dense waters. Internal waves are generated by three main mechanisms: Moon- and Sun-generated tidal flow over steep or rough topography (lower right); fluctuating wind stress on the ocean surface (upper left); and quasi-steady flow over rough topography (lower left). This last mechanism, which produces the internal lee waves investigated by Nikurashin and Ferrari<sup>2</sup>, is analogous to mountain waves in the atmosphere.



the meridional overturning circulation<sup>6</sup>.

Their theoretical argument for the global importance of lee-wave breaking is provocative and intriguing. However, the quantitative part of their conclusion is subject to several uncertainties. The largest issue is the paucity of direct observations of lee-wave generation, propagation and turbulent breaking in the deep or abyssal ocean. As a starting point, the authors suggest that internal lee waves have similar vertical propagation and decay properties to internal tides, which have yet to be observationally explored. Data collected as part of the Diapycnal and Isopycnal Mixing Experiment in the Southern Ocean, analysis of which is just beginning, may provide the first such evidence (see <http://dimes.ucsd.edu>).

By comparison, a decade of focused experiments on internal-tide processes has mapped out the large range of wave generation and breaking behaviour, and has allowed differentiation between contending theories<sup>7</sup>. Results from those experiments and closely integrated theoretical work are now being turned into new parameterizations of turbulent mixing for use in global climate models (see [go.nature.com/vylsba](http://go.nature.com/vylsba)). Nikurashin and Ferrari make a persuasive argument that internal lee waves may have such a substantial role in deep-ocean mixing that they should be included in future climate models. The next step towards that goal should be targeted deep-ocean observational studies, because the ocean never fails to surprise. ■

**Jennifer MacKinnon** is at the Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California 92093, USA.

e-mail: [jmackinn@ucsd.edu](mailto:jmackinn@ucsd.edu)

1. Schiermeier, Q. *Nature* **447**, 522–524 (2007).
2. Nikurashin, M. & Ferrari, R. *Geophys. Res. Lett.* **40**, 3133–3137 (2013).
3. Alford, M. H. et al. *J. Phys. Oceanogr.* **41**, 2211–2222 (2011).
4. MacKinnon, J., Laurent, L. S. & Garabato, A. N. in *Ocean Circulation and Climate: A 21st Century Perspective* 2nd edn (eds Siedler, G., Griffies, S., Gould, J. & Church, J.) (Academic, 2013).
5. Melet, A., Hallberg, R., Legg, S. & Polzin, K. J. *J. Phys. Oceanogr.* **43**, 602–615 (2013).
6. Talley, L. D. *Oceanography* **26**, 80–97 (2013).
7. Polzin, K. L. *Ocean Modelling* **30**, 298–309 (2009).

peptide, stiffened lipid-bilayer membranes into which it was incorporated in a manner akin to the action of cholesterol. They propose that stiffening of the trypanosome lysosomal membrane by TgsGP prevents APOL1 membrane insertion. This concept remains to be verified, but it opens up the possibility that cholesterol-rich membranes may be resistant to APOL1 and inspires the question of whether cholesterol protects human cells from APOL1-mediated injury.

However, Uzureau and colleagues also show that introducing the *TgsGP* gene into a non-resistant trypanosome species does not confer resistance. Further studies of the TLF-uptake pathway suggested that resistant parasites have, in addition to TgsGP, altered lysosomal pH and protein-cleaving enzymatic activity that enhances their resistance to TLF.

*T. b. gambiense* type 1 is also known to have evolved variations in the TLF receptor, one of which is a single amino-acid change that is found in all strains of this subspecies<sup>5</sup>. This modification reduces the affinity of the receptor for its ligand<sup>6</sup>, and thereby reduces the amount of TLF accumulated by the parasite<sup>7</sup>. Uzureau *et al.* speculate that this receptor variant arose during evolution as a result of selection pressure from humans with chronic low levels of red-blood-cell lysis owing to diseases such as malaria. Lysis of red blood cells releases haemoglobin, which binds to the protein haptoglobin, an abundant component of serum. This results in the clearance of haptoglobin by immune cells, causing a condition known as hypohaptoglobinaemia. The released haemoglobin also binds to HPR on TLF and this, coupled with the removal of competing haptoglobin–haemoglobin complexes, results in increased TLF uptake by trypanosomes, leading to enhanced parasite killing. Uzureau *et al.* show that the mutation in the *T. b. gambiense* receptor, which reduces TLF uptake, allows parasite survival in hypohaptoglobinaemic serum.

Previous work from the same research group<sup>3</sup> had shown that the East African

## PARASITOLOGY

# Molecular one-upmanship

**Pathogens and their hosts engage in perpetual molecular arms races. In one such evolutionary stand-off, the protagonists are trypanosome parasites and a human immune complex based on a high-density lipoprotein. SEE LETTER P.430**

JAYNE RAPER & DAVID J. FRIEDMAN

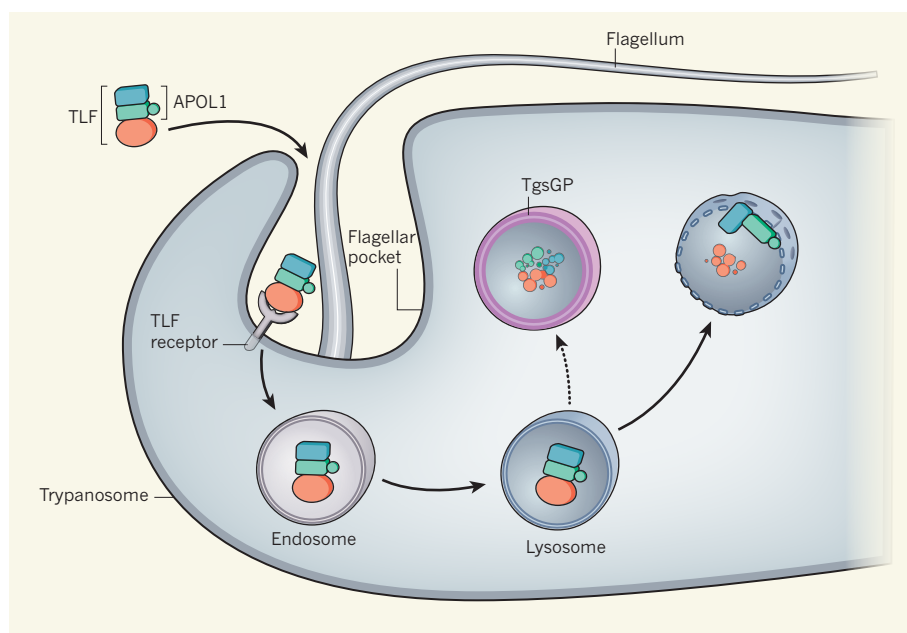
**A**frican trypanosomes, the cause of African sleeping sickness, have waged war against humans for millennia. The epicentre of this struggle is the rapidly evolving protein apolipoprotein L1, which is found in humans and some other primates<sup>1</sup>, and is the key mediator of these hosts' innate immune responses to the parasites. But some trypanosomes have evolved to overcome this defence and can successfully establish long-term infections in humans. On page 430 of this issue, Uzureau *et al.*<sup>2</sup> describe the complex suite of resistance adaptations used by the most prevalent disease-causing African trypanosome, *Trypanosoma brucei gambiense*.

African trypanosomes are single-celled parasites that live freely in their host's bloodstream. To escape immune destruction, they constantly change their surface coat by selecting from a repertoire of genes that encode variant surface glycoproteins. But some primates, including humans, have developed innate-immune complexes known as trypanolytic factors (TLFs) that can circumvent this antigenic variation and kill most species of trypanosome. TLFs are high-density lipoprotein molecules that have two key protein components: haptoglobin-related protein (HPR) and apolipoprotein L1 (APOL1). HPR binds to the protein haemoglobin and this complex binds to a receptor on trypanosomes. The trypanosomes

then take up the TLF and transfer it to a digestive cellular organelle called the lysosome in order to use its haemoglobin and lipids for their own biosynthetic processes. But during this transit, APOL1 is released and inserted into the lysosome membrane, where it forms small ion-permeable pores that cause the parasite to burst and die<sup>3</sup> (Fig. 1).

However, some human-infective trypanosomes, including *T. b. gambiense*, which causes 97% of human cases of sleeping sickness, can resist the activity of TLFs. By combining previous observations with new findings, Uzureau *et al.* propose that the parasites have three requirements for complete TLF resistance: the expression of *T. b. gambiense*-specific glycoprotein (TgsGP); modification of the TLF receptor; and changes in lysosomal physiology.

The protein TgsGP is unique to the subspecies *T. b. gambiense* type 1 (ref. 4) and is used as a marker for these parasites. Deletion of the *TgsGP* gene renders these trypanosomes sensitive to human TLFs, and adding back the gene restores resistance<sup>2</sup>. To determine which domain of the protein is responsible for TLF resistance, Uzureau *et al.* introduced a battery of modified *TgsGP* genes into *TgsGP*-deficient parasites. Despite a limitation of this approach — the potential for misfolding of the modified TgsGPs that interferes with their activity — the authors identified a particular stretch of amino-acid residues that was required for resistance and that, when synthesized as a



**Figure 1 | Trypanosome resistance to host immunity.** The trypanolytic factors (TLFs) found in human serum are high-density lipoprotein molecules that include the protein APOL1. TLFs bind to receptors on the trypanosome surface within the flagellar pocket. The TLFs are then internalized and transported through intracellular organelles for degradation — first to endosomes and then to the lysosome. The acidic pH of the lysosome induces a change in the conformation of APOL1 such that it is released from the TLF and incorporated into the lysosome membrane. This leads to the formation of membrane pores that upset the intracellular ionic balance and cause the trypanosome to burst and die. Uzureau *et al.*<sup>2</sup> propose that *Trypanosoma brucei gambiense* type 1 combines three processes to resist this immune defence: the expression of the protein TgsGP, which stiffens the lysosomal membrane and prevents APOL1 incorporation, modification of the TLF receptor and changes in lysosomal physiology.

human-infective trypanosome *T. b. rhodesiense* has evolved a different way to escape killing by TLF. This species has not acquired mutations in the TLF receptor, even though malaria (and therefore hypohaptoglobinaemia) is endemic throughout sub-Saharan Africa. Instead, the parasite has evolved a protein called the serum-resistance-associated protein (SRA)<sup>8</sup>, which alone is sufficient to confer complete resistance to the TLFs present in most humans. SRA binds to APOL1 with high affinity and prevents either the insertion and/or oligomerization of APOL1 in membranes<sup>9</sup>. However, some people produce APOL1 variants (called G1 and G2) that arose in Africans after the migration of ancestral humans out of Africa; sera from individuals with these variant proteins kill *T. b. rhodesiense* *in vitro*<sup>10</sup>, but their ability to protect against infection has yet to be confirmed. There are multiple African-restricted APOL1 variants in addition to G1 and G2, and one of these may prove to confer resistance to *T. b. gambiense*<sup>11</sup>.

One surprising consequence of these evolutionary events is that African-Americans who express only the G1 or G2 APOL1 variants have 7–30-fold higher rates of several types of kidney disease<sup>10,12</sup>. This scenario resembles the human mutations in haemoglobin that provide protection against malaria but lead to the blood disorder sickle-cell anaemia. It is possible that kidney disease or other deleterious effects of

these trypanosome-protective APOL1 variants may have prevented them from becoming fixed throughout the population.

Uzureau *et al.* have contributed a significant advance to our understanding of natural selection involving African trypanosomes and humans, but their work points towards ever more complex and interrelated events involving different trypanosome species, and

perhaps malaria parasites, in this long-running battle with humans. So who is winning this epic arms race? Perhaps it is the non-human primates — baboons, for example, express APOL1 variants that seem capable of killing human-infective trypanosomes<sup>13,14</sup> but that do not cause kidney disease. Figuring out the secrets of baboon APOL1 may answer crucial questions that will help to solve the riddles of both African sleeping sickness and kidney disease. ■

Jayne Raper is at Hunter College, City University of New York, New York, New York 10065, USA. David J. Friedman is in the Renal Division, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts 02215, USA. e-mails: raper@genectr.hunter.cuny.edu; dfriedma@bidmc.harvard.edu

- Smith, E. E. & Malik, H. S. *Genome Res.* **19**, 850–858 (2009).
- Uzureau, P. *et al. Nature* **501**, 430–434 (2013).
- Vanhollebeke, B. & Pays, E. *Mol. Microbiol.* **76**, 806–814 (2010).
- Gibson, W., Nemetschke, L. & Ndung'u, J. *Infect. Genet. Evol.* **10**, 453–458 (2010).
- Symula, R. E. *et al. PLoS Negl. Trop. Dis.* **6**, e1728 (2012).
- Higgins, M. K. *et al. Proc. Natl Acad. Sci. USA* **110**, 1905–1910 (2013).
- Kieft, R. *et al. Proc. Natl Acad. Sci. USA* **107**, 16137–16141 (2010).
- De Greef, C., Chimfwembe, E., Kihang'a Wabacha, J., Bajjya Songa, E. & Hamers, R. *Ann. Soc. Belg. Med. Trop.* **72**, Suppl. 1, 13–21 (1992).
- Stephens, N. A., Kieft, R., MacLeod, A. & Hajduk, S. L. *Trends Parasitol.* **28**, 539–545 (2012).
- Genovese, G. *et al. Science* **329**, 841–845 (2010).
- Ko, W.-Y. *et al. Am. J. Hum. Genet.* **93**, 54–66 (2013).
- Kopp, J. B. *et al. J. Am. Soc. Nephrol.* **22**, 2129–2137 (2011).
- Thomson, R., Molina-Portela, P., Mott, H., Carrington, M. & Raper, J. *Proc. Natl Acad. Sci. USA* **106**, 19509–19514 (2009).
- Kageruka, P. *et al. Ann. Soc. Belg. Med. Trop.* **71**, 39–46 (1991).

#### MATERIALS SCIENCE

## Fast-track solar cells

The ultimate goal of the solar-cell industry is to make inexpensive devices that are highly efficient at converting sunlight into electricity. The advent of perovskite semiconductors could be the key to reaching this goal. [SEE LETTER P.395](#)

MICHAEL D. MCGEEHEE

After the discovery of a new semiconductor that has potential for solar-cell applications, it has always taken numerous teams of researchers more than ten years to figure out how to use it to make devices that convert sunlight into power with greater than 15% efficiency. Liu and colleagues' report<sup>1</sup> on page 395 of this issue is therefore

truly remarkable\*. They describe cells that have slightly greater than 15% conversion efficiency, and which are based on perovskites, a class of semiconductor material that was first used for this application just four years ago. This follows hot on the heels of another *Nature* paper by Burschka *et al.*<sup>2</sup>, who in July reported a perovskite-containing solar cell that reached the 15%

\*This article and the paper under discussion<sup>1</sup> were published online on 11 September 2013.





**Figure 1 | Solar-power generator in Tokyo.** Electricity generated from sunlight is already economically viable in many regions. Solar cells containing perovskite materials, such as those reported in two recent papers<sup>1,2</sup>, might one day convert sunlight into power more efficiently than currently available devices.

conversion milestone. The fact that multiple teams are making such rapid progress suggests that the perovskites have extraordinary potential, and might elevate the solar-cell industry to new heights.

Researchers started using organometal trihalide perovskite semiconductors (which have the formula  $(\text{CH}_3\text{NH}_3)\text{PbX}_3$ , where Pb is lead and X can be iodine, bromine or chlorine) in solar cells in 2009 (ref. 3). Since then, the performance of perovskite-containing solar cells has skyrocketed.

Initially, perovskites were used as dye replacements in dye-sensitized solar cells. In these devices, dyes act as light absorbers coating the surface of a film of titanium dioxide ( $\text{TiO}_2$ ) nanoparticles. When light is absorbed by a dye, electrons and positive-charge carriers known as holes are generated and transferred to different transport materials: to  $\text{TiO}_2$  for electrons, and to another material for holes. The transport materials then carry the charges to separate electrodes and so generate a voltage. Burschka and co-workers raised the world record for conversion efficiency in this kind of solar cell from 12% by packing lots of light-absorbing perovskite into the  $\text{TiO}_2$  film, so that most of the light striking the cell was absorbed in a very thin film.

Liu *et al.* have now shown that the perovskite semiconductor not only strongly absorbs light, but also transports both positive and negative charges. This allowed them to do away with the  $\text{TiO}_2$  nanoparticles, and so to implement a more conventional solar-cell architecture

than that of Burschka and co-workers: in Liu and colleagues' device, the light-absorbing semiconductor is sandwiched between electron- and hole-selective electrodes. Remarkably, it maintains high conversion efficiency.

The rapid emergence of perovskites comes at an interesting time in the solar-cell industry's history. In locations where sunlight is abundant and electricity is fairly expensive, the production of electricity using solar cells is now cost-competitive with that of conventional electricity sources (Fig. 1). However, there is still a need to reduce the costs of making and installing solar cells, and to improve conversion efficiency so that fewer panels need to be installed.

Almost all solar cells manufactured today are made with approximately 150-micrometre-thick wafers of crystalline silicon, and have conversion efficiencies in the range of 17–23%. For decades, many researchers have attempted to develop high-efficiency and lower-cost alternatives to crystalline silicon by depositing semiconductor films that are less than a few micrometres thick on inexpensive substrates such as glass, metal or plastic. Of the thousands of semiconductor materials that have been tried, only a few (such as cadmium telluride and copper indium gallium selenide) have enabled conversion efficiencies in the range of 15–20% (ref. 4).

Moreover, the efficiencies of such cells are limited by the many defects formed during the rapid deposition of the semiconductor films.

These defects promote the recombination of electrons and holes, which drains power and reduces voltage. Because thin-film solar cells are not as efficient as crystalline silicon cells, they are less desirable to customers, who want to minimize installation costs, and so most of the companies making them are going out of business.

It is therefore scientifically interesting and technologically important that the voltage produced by the perovskite solar cells<sup>1,2</sup> is more than 1 volt; silicon and most thin-film solar cells typically produce only 0.7 V under 'open circuit' conditions. Apparently, there is something special about perovskites that slows down recombination. What this special something is remains unknown, making it hard to estimate how efficient the perovskite solar cells could be, but it seems likely that their efficiency will continue to climb. It is even possible that perovskites will emerge as the champion material for solar cells.

Perovskites and crystalline silicon need not compete with each other as solar-cell materials. Because silicon has a smaller bandgap (the energy required to generate conducting electrons) than the perovskites, it absorbs a chunk of the solar spectrum that perovskites do not. One could therefore place a silicon cell underneath a perovskite cell to form a tandem cell<sup>5</sup>. Perovskites could be printed on top of silicon in a way that would add little to the manufacturing cost, and because perovskites generate a higher voltage than silicon, the tandem cell would be more efficient than a silicon cell. A rough calculation suggests that the performance of currently available, commercial solar cells could be boosted by 25%, and would require only minor modifications to existing factories.

The news that perovskites perform extraordinarily well in solar cells has spread fast, and many academic and industrial researchers are already discussing their potential for commercialization. Most are optimistic that the efficiencies of perovskite cells will continue to improve until they are competitive with, or better than, existing commercially available solar cells. However, there are concerns that lead toxicity will be a problem, because the perovskites are soluble in water and could be washed out of a leaky module. The race is on to discover perovskites containing non-toxic elements that have the same desirable properties as the lead compounds.

Another possible issue is the long-term stability of perovskites. The results of preliminary tests performed on short timescales are promising, but some researchers remain concerned that a water-soluble material that can be sublimed at a low temperature (as perovskites can) will lack the rock-like stability that enables silicon solar cells to last for more than 25 years.

Further breakthroughs in perovskite solar cells can certainly be expected in the months

and years to come. The history of these materials may be brief, but their performance so far suggests that their role in the solar-cell field will be anything but. ■

**Michael D. McGehee** is in the Department of Materials Science and Engineering, Stanford University, Stanford, California 94305-4045, USA.

e-mail: mmcgehee@stanford.edu

1. Liu, M., Johnston, M. B. & Snaith, H. J. *Nature* **501**, 395–398 (2013).
2. Burschka, J. *et al.* *Nature* **499**, 316–319 (2013).
3. Kojima, A., Teshima, K., Shirai, Y. & Miyasaka, T. *J. Am. Chem. Soc.* **131**, 6050–6051 (2009).
4. [www.nrel.gov/ncpv](http://www.nrel.gov/ncpv)
5. Bailey, Z. M. & McGehee, M. D. *Energ. Environ. Sci.* **5**, 9173–9179 (2012).

## STEM CELLS

# Down's syndrome link to ageing

**Triplication of the enzyme USP16 in models of Down's syndrome creates defects in the stem cells resident in adult tissues. This finding provides insight into stem-cell homeostasis during ageing. [SEE ARTICLE P.380](#)**

**GEORGE P. SOUROULLAS  
& NORMAN E. SHARPLESS**

**D**own's syndrome is a complex developmental disorder involving congenital malformation and mental retardation. It affects 1 in about 700 individuals, and is caused by an extra copy of chromosome 21. Individuals with Down's syndrome therefore have three copies (trisomy) of some 300 protein-encoding genes, and of an even larger number of non-coding RNA sequences. Although a few candidate genes have been linked to the spectrum of disorders associated with Down's syndrome, generally it is unclear how trisomy of specific genes contributes to the condition. In a paper published in this issue, Adorno *et al.*<sup>1</sup> (page 380) implicate the deubiquitinating enzyme Usp16 in an impaired ability of adult-tissue stem cells to self-renew\*.

As childhood death among individuals with Down's syndrome has declined sharply, it has become clear that the condition is associated with an early onset of age-related disorders, such as Alzheimer's disease, immune dysfunction, premature menopause and disorders of the endocrine system<sup>2</sup>. Consequently, gerontologists have long wondered whether Down's syndrome is a form of segmental progeria<sup>3</sup> — that is, whether its diverse aspects result from a shared mechanism with accelerated ageing (progeria).

But Down's syndrome is distinct from progerias that result from alterations in DNA metabolism (such as Werner's and Bloom syndromes and dyskeratosis congenita), because cells from individuals with Down's syndrome do not show significant defects

in DNA metabolism or repair. Moreover, affected individuals exhibit a peculiar pattern of cancer incidence — an increased risk of childhood disorders associated with the proliferation of haematopoietic (blood-lineage) cells, but a much reduced risk of more common solid tumours<sup>4,5</sup>. By contrast, DNA-metabolism disorders associated with progeria are accompanied by a general increase in cancer incidence.

Against this confusing backdrop, Adorno *et al.* suggest an explanation for accelerated ageing in Down's syndrome in terms of impaired function of somatic stem cells — tissue-resident cells that self-renew throughout life to repair damaged cells; the functional decline of these stem cells is thought to underlie aspects of mammalian ageing. Using two mouse models of Down's syndrome, as well as cultured human cells, the authors show that certain aspects of Down's syndrome, including phenotypes (traits) of premature ageing, are associated with misregulation of the machinery involved in cellular senescence — a state of permanent growth arrest associated with tumour suppression and ageing.

Of the two mouse models, one (Ts65Dn) carries three copies of around two-thirds of the 132 genes that are the equivalent of those triplicated on human chromosome 21, whereas the other is trisomic for 79 of these 132 genes. The authors examined the self-renewal capacity of stem cells and progenitor cells obtained from the bone marrow and brain and breast tissue of these animals. They observed disrupted homeostasis of somatic stem cells in Ts65Dn mice, and found that cultured cells isolated from these mice showed signs of increased senescence. A comparison of the triplicated chromosomal regions in the two models led the researchers to suspect trisomy of the *Usp16*



## 50 Years Ago

Sir Nevill Mott spoke about research in the universities and repudiated the view that university research is just an expensive luxury to stimulate the dons and keep them up to date ... The third and last discussion ranged widest of all. There were those who took for granted, so it seemed, that the modern undergraduate at entry was fundamentally perverse and irresponsible, overdeveloped intellectually and socially as well as emotionally immature; those who defended him warmly against the criticisms of his elders; those who felt that he ... was too much 'of the world' already and needed to be withdrawn during his university years from the pressure and hurly-burly of ordinary existence; for them the hall of residence on the university campus was the right answer; those, on the contrary, who felt that segregation — whether from family or other social ties — and an artificially high standard of living were wholly bad; these approved of the civic university whose students live with their own families ... But all united in one aim — to prevent, if possible, the university from turning out a lop-sided individual, intellectually precocious but humanly underdeveloped.

**From *Nature* 21 September 1963**

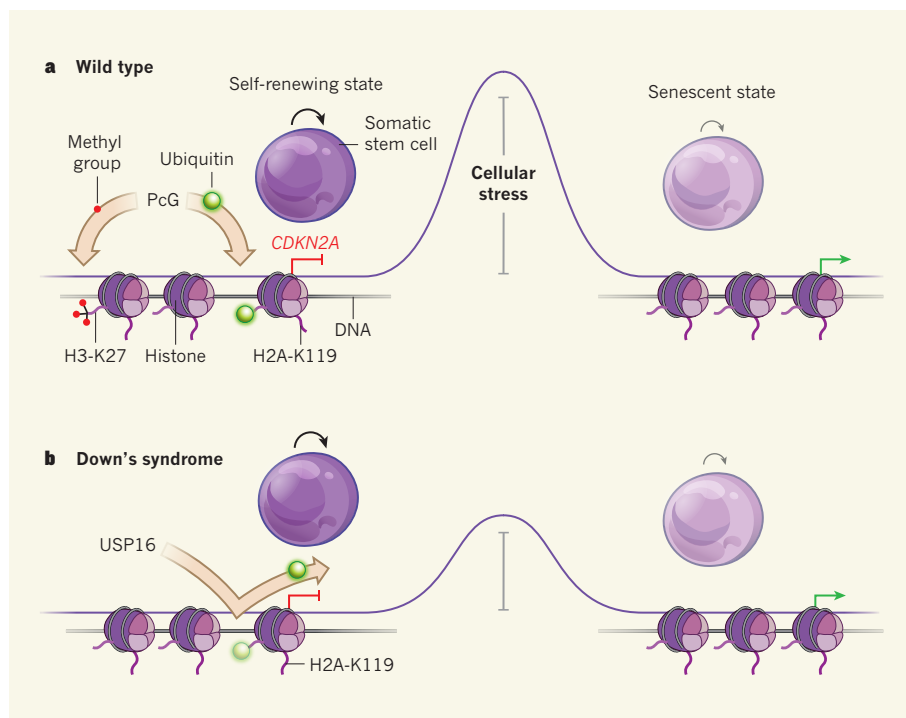
## 100 Years Ago

A curious story comes from Ireland that Mr. E. S. Dodgson, of Jesus College, Oxford, has discovered at Killult, Falcarragh, Donegal, a stone said to contain an Ogham inscription giving a clue to a great treasure concealed in the neighbourhood by an ancient Irish chieftain ... We wish the discoverer success in unearthing the treasure, but until he succeeds, or some other interpretation of the supposed inscription is suggested, it may be well to reserve opinion on the matter.

**From *Nature* 18 September 1913**

\*This article and the paper under discussion<sup>1</sup> were published online on 11 September 2013.





**Figure 1 | An altered threshold for stress.** **a**, Normally, Polycomb group (PcG) proteins chemically modify DNA-associated histone proteins to regulate gene expression. To support self-renewal of somatic stem cells, PcG proteins suppress the expression of *CDKN2A* and other genomic loci through specific biochemical modifications (addition of trimethyl groups to the H3-K27 amino-acid residue and of the ubiquitin protein to H2A-K119). Extended exposure to cellular stress (such as DNA damage) can cause self-renewing cells to enter a senescent state, through loss of PcG-mediated histone modifications, leading to derepression of the associated genomic loci. **b**, The threshold for conversion between these states can determine the rate of molecular ageing. Adorno *et al.*<sup>1</sup> suggest that, in Down's syndrome, *USP16* trisomy reduces H2A-K119 ubiquitination in certain self-renewing tissues, thereby reducing this threshold and promoting somatic-stem-cell senescence.

gene as a potential cause of reduced stem-cell self-renewal.

This finding is intriguing, because USP16 is thought to counteract gene silencing mediated by the Polycomb group (PcG) of proteins. Specifically, this enzyme detaches a ubiquitin protein from H2A-K119 (the lysine amino-acid residue at position 119 of the DNA-associated histone protein H2A). Because PcG-mediated gene repression is crucial for somatic-stem-cell function, Adorno *et al.* hypothesized that trisomy of *USP16* might adversely alter the epigenetic state of these cells. (Epigenetic modifications describe alterations, such as ubiquitination, to DNA or histone proteins that regulate gene expression without changing the DNA sequence.)

One genomic region (locus) that is tightly repressed in a PcG-dependent manner in healthy cells is *CDKN2A*; stress, however, attenuates this PcG-mediated repression (Fig. 1a). Activation of *CDKN2A*, which encodes two tumour-suppressor proteins, p16<sup>Ink4a</sup> and p19<sup>Arf</sup>, in turn promotes cellular senescence. Accordingly, Adorno and colleagues demonstrate that reduced USP16 expression in Down's syndrome cells leads to repression of *CDKN2A* and recovery from self-renewal defects. So it seems that

*USP16* trisomy leads to a 'Polycomb-opathy' in Down's syndrome, whereby diminished repression of *CDKN2A*, and perhaps of other ageing-associated loci, contributes to premature senescence of somatic stem cells (Fig. 1b).

These findings are exciting, but come with certain caveats. For one, it is unlikely that a single gene accounts for the disparate phenotypes of Down's syndrome, and even for the subset of ageing-like phenotypes. Trisomy of the *APP* gene, for example, has also been linked to increased incidence of Alzheimer's disease in Down's syndrome<sup>6</sup>.

Also, despite significant interest in haematopoietic stem cells (HSCs) in Down's syndrome, proliferative defects have not been reported in HSCs of humans with the disorder. This discrepancy may reflect technical issues in purifying human HSCs, which could make their functional characterization difficult. Moreover, trisomy of the gene *ERG* predisposes individuals with Down's syndrome to disorders associated with proliferation of cells in the bone marrow<sup>7</sup>, further complicating HSC analysis. Technical differences and predilection to such disorders might also explain why an earlier study in Ts65Dn mice reported<sup>8</sup> an expanded pool of haematopoietic progenitor cells.

These reservations notwithstanding, the

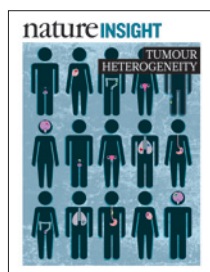
authors' study has compelling features. First, this work and another recent study<sup>9</sup> suggest that Polycomb-opathy is an emerging theme in segmental progeria. The other study reported pronounced, genome-wide alterations in a related histone mark — trimethylation of lysine 27 of histone 3 (H3-K27) — in fibroblast cells from individuals with the most severe form of progeria, Hutchinson–Gilford syndrome. Like H2A-K119 ubiquitination, this addition of three methyl groups to histone 3 is catalysed by PcG complexes (Fig. 1a). And these enzymatic activities function in concert to effect heritable repression of *CDKN2A* and other loci.

Furthermore, unbiased genome-wide association studies have linked the *CDKN2A* locus to diverse age-associated phenotypes such as atherosclerosis, type 2 diabetes, glaucoma and susceptibility to cancer<sup>10</sup>. Along with related mouse studies<sup>11,12</sup>, these results hint at the phenotypic impact of excess *CDKN2A* expression in humans — reduced susceptibility to cancer and atherosclerosis, and increased immunodeficiency. This is a reasonable description of ageing in Down's syndrome, which is characterized by an increased risk of death from infection, but decreased mortality from cancer and atherosclerosis<sup>4,5</sup>.

Another appealing aspect of Adorno and colleagues' work is the suggestion that the USP16 enzyme could be a target for treating some aspects of Down's syndrome. In this regard, however, it is important to temper any optimism, because even if selective USP16 inhibitors were developed, clinical trials for age-related conditions in Down's syndrome would probably be challenging. Nevertheless, in suggesting that USP16 trisomy creates a congenital disruption of the PcG-regulated epigenetic state in somatic stem cells, the present work furthers our understanding of stem-cell homeostasis during ageing. ■

**George P. Souroullas and Norman E. Sharpless** are in the Departments of Medicine and Genetics, Lineberger Comprehensive Cancer Center, University of North Carolina School of Medicine, Chapel Hill, North Carolina 27599-7295, USA.  
e-mail: norman\_sharpless@med.unc.edu

- Adorno, M. *et al.* *Nature* **501**, 380–384 (2013).
- Esbensen, A. J. *Int. Rev. Res. Ment. Retard.* **39**, 107–126 (2010).
- Martin, G. M. *Fed. Proc.* **38**, 1962–1967 (1979).
- Yang, Q., Rasmussen, S. A. & Friedman, J. M. *Lancet* **359**, 1019–1025 (2002).
- Englund, A., Jonsson, B., Zander, C. S., Gustafsson, J. & Annerén, G. *Am. J. Med. Genet. A* **161**, 642–649 (2013).
- Salehi, A. *et al.* *Neuron* **51**, 29–42 (2006).
- Ng, A. P. *et al.* *Blood* **115**, 3966–3969 (2010).
- Kirsammer, G. *et al.* *Blood* **111**, 767–775 (2008).
- McCord, R. P. *et al.* *Genome Res.* **23**, 260–269 (2013).
- Jeck, W. R., Siebold, A. P. & Sharpless, N. E. *Ageing Cell* **11**, 727–731 (2012).
- Liu, Y. *et al.* *Blood* **117**, 3257–3267 (2011).
- Kuo, C. L. *et al.* *Arterioscler. Thromb. Vasc. Biol.* **31**, 2483–2492 (2011).



**Cover illustration**  
Nik Spencer

**Editor, *Nature***  
Philip Campbell

**Publishing**  
Richard Hughes

**Production Editor**  
Jenny Rooke

**Art Editor**  
Nik Spencer

**Sponsorship**  
David Bagshaw  
Yvette Smith  
Reya Silao

**Production**  
Ian Pope

**Marketing**  
Elena Woodstock  
Steven Hurst

**Editorial Assistant**  
Abbie Williams

The Macmillan Building  
4 Crinan Street  
London N1 9XW, UK  
Tel: +44 (0) 20 7833 4000  
e: [nature@nature.com](mailto:nature@nature.com)



nature publishing group

Cancer is not one but many diseases. It is different in each patient and continuously evolves into a progressively complex interplay of diverse tumour cells with their changing environment.

In an insightful and prescient 1976 paper, Peter Nowell proposed the now prevailing view of cancer as a process of clonal evolution, in which successive rounds of clonal selection give rise to tumours with diverse genetic and other molecular alterations that may necessitate ‘personalized’ therapies. Functional and phenotypic heterogeneity of cancer cells has also long been recognized, with experiments as early as the 1930s demonstrating that some but not other cells from mouse tumours could give rise to new tumours when transplanted.

The picture that has emerged over the past few decades — especially with the advent of more sophisticated model systems and technologies — is of even greater than anticipated genetic, phenotypic and functional heterogeneity and plasticity within tumours and between primary tumours and metastases. Adding to this bewildering complexity is the heterogeneity of the tumour micro-environment, inflammatory stimuli, the immune response, mechanical stresses, therapeutic intervention and many other factors, such as diet and the microbiota. These continuously changing environmental influences affect which cancer cell subpopulations are able to survive, proliferate, spread and resist therapy.

If every cancer, and perhaps every cancer cell, is unique, and some cancer cell populations are more ‘unique’ than others, this needs to be taken into consideration for the improvement of cancer diagnoses and prognoses, for the treatment and monitoring of each patient and for the design of clinical trials to evaluate new therapies.

This Insight covers some of the key topics relevant to the fast-moving field of cancer research, ranging from the basic understanding of tumour heterogeneity to translational cancer research and clinical trials. We hope to have captured some of the most exciting recent developments in this area.

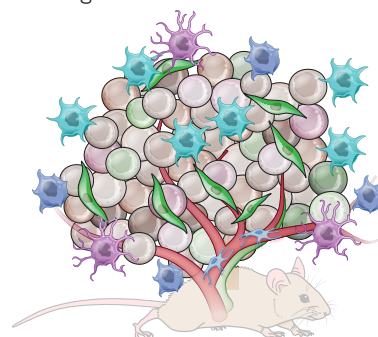
*Nature* is pleased to acknowledge the financial support of Gilead Sciences in producing this Insight. As always, *Nature* carries sole responsibility for all editorial content.

**Barbara Marte**  
*Senior Editor*

### CONTENTS

#### REVIEWS

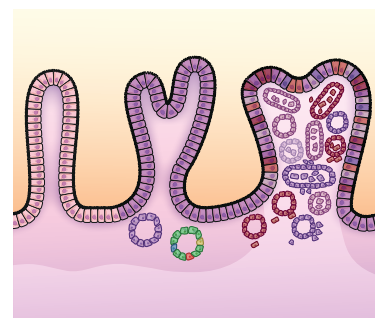
- 328 Tumour heterogeneity and cancer cell plasticity**  
*Corbin E. Meacham & Sean J. Morrison*
- 338 The causes and consequences of genetic heterogeneity in cancer evolution**  
*Rebecca A. Burrell, Nicholas McGranahan, Jiri Bartek & Charles Swanton*
- 346 Influence of tumour micro-environment heterogeneity on therapeutic response**  
*Melissa R. Junttila & Frederic J. de Sauvage*



- 355 Tumour heterogeneity in the clinic**  
*Philippe L. Bedard, Aaron R. Hansen, Mark J. Ratain & Lillian L. Siu*

#### PERSPECTIVE

- 365 Selection and adaptation during metastatic cancer progression**  
*Christoph A. Klein*





# Tumour heterogeneity and cancer cell plasticity

Corbin E. Meacham<sup>1</sup> & Sean J. Morrison<sup>1</sup>

**Phenotypic and functional heterogeneity arise among cancer cells within the same tumour as a consequence of genetic change, environmental differences and reversible changes in cell properties. Some cancers also contain a hierarchy in which tumorigenic cancer stem cells differentiate into non-tumorigenic progeny. However, it remains unclear what fraction of cancers follow the stem-cell model and what clinical behaviours the model explains. Studies using lineage tracing and deep sequencing could have implications for the cancer stem-cell model and may help to determine the extent to which it accounts for therapy resistance and disease progression.**

The cancer stem-cell model provides one explanation for the phenotypic and functional heterogeneity among cancer cells in some tumours<sup>1–5</sup>. The model posits that some cancers are organized into a hierarchy of subpopulations of tumorigenic cancer stem cells and their non-tumorigenic progeny. In these cases, cancer stem cells are thought to drive tumour growth and disease progression, perhaps through therapy resistance<sup>6–8</sup> and metastasis<sup>9,10</sup>. However, difficulty replicating solid-cancer stem-cell markers, variability from patient to patient and variation in results from different xenograft models have meant that it is unclear what fraction of cancers follow this model — the majority or only a minority<sup>11</sup>?

Even in cancers that clearly contain a hierarchy of tumorigenic and non-tumorigenic cells, this hierarchy must coexist with other sources of heterogeneity, including clonal evolution<sup>12</sup> (see the Review by Swanton and colleagues on page 338), heterogeneity in the micro-environment<sup>13,14</sup> (see the Review by Junttila and de Sauvage on page 346) and reversible changes in cancer-cell properties that can occur independently of hierarchical organization<sup>15–18</sup>. In these circumstances it is not necessarily clear which phenotypic and functional differences among cells arise from which sources of heterogeneity. To what extent do metastasis, therapy resistance and disease progression reflect the intrinsic properties of cancer stem cells as opposed to genetic evolution or other sources of heterogeneity? Integration of results from multiple experimental approaches will be necessary to distinguish the relative contributions of these sources of heterogeneity to disease progression.

New experimental approaches have provided perspective and insight into this question. Genetic approaches to fate map the contributions of cancer cells to tumour growth in mice have provided evidence to support the cancer stem-cell model in some contexts, and evidence against the model in others<sup>19–23</sup>. Transplantation assays evaluate the potential of cancer cells to form tumours, rather than their actual fate in the native tumour, and so fate mapping complements what we have learned from transplantation assays (Fig. 1). High-coverage sequencing of human tumours has also provided insight into genetic heterogeneity within tumours and the cells that are responsible for relapse after therapy<sup>24–28</sup>. In this Review, we evaluate the implications of these new data for the cancer stem-cell model and the extent to which this model accounts for clinically important forms of heterogeneity in cancer.

## Tests for tumorigenic potential

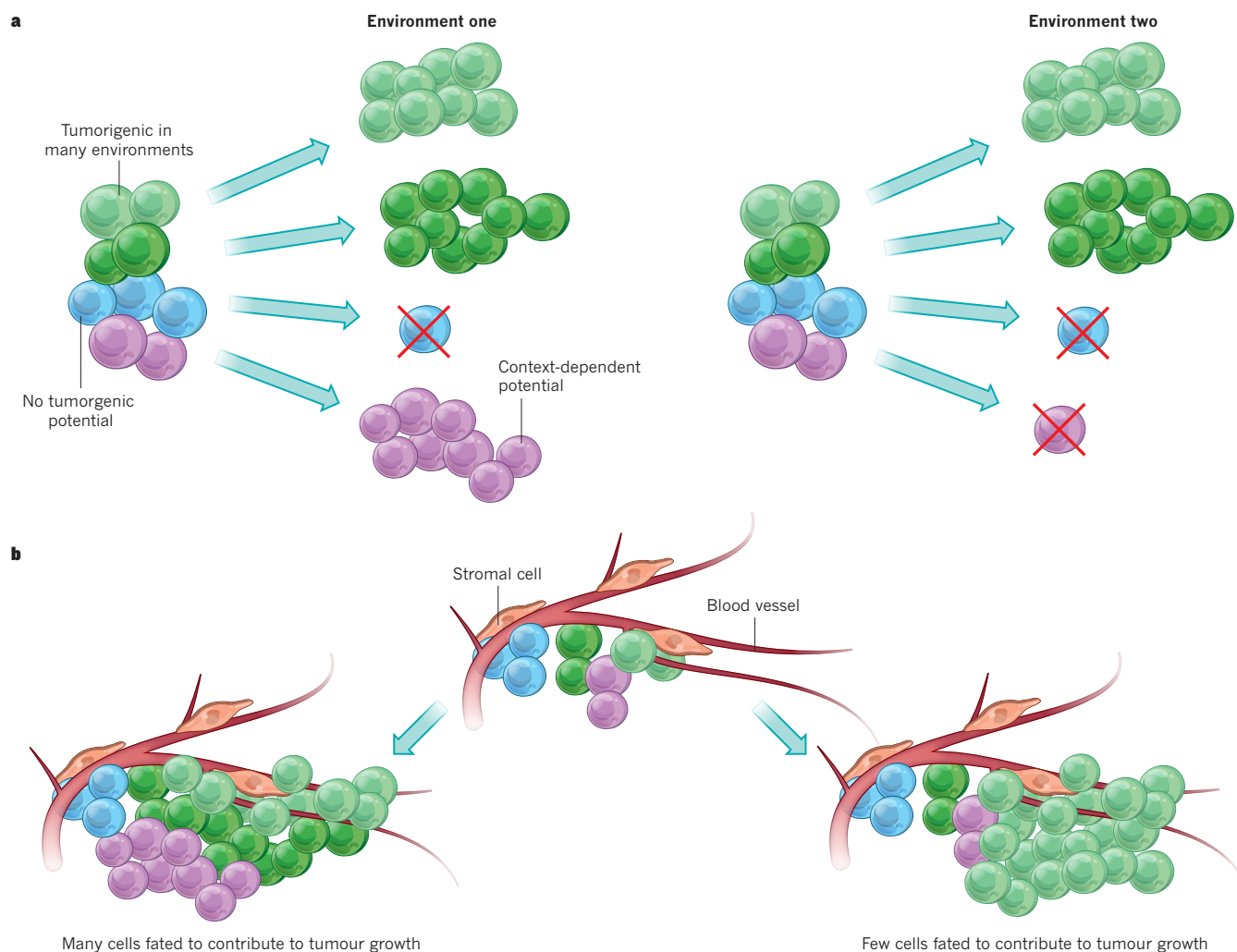
Central to the cancer stem-cell model is the idea that tumour growth and disease progression are driven by minority populations of tumorigenic

cells, and that most other cancer cells have little or no capacity to contribute to tumour growth. This means that therapeutic strategies should focus particularly on killing tumorigenic cells. In experiments, the cancer stem-cell model has mainly been tested using transplantation assays, which test the potential of a cancer cell to form a tumour. These assays have demonstrated the existence of phenotypically distinct subpopulations of tumorigenic and non-tumorigenic cells in a number of human cancers, including acute myeloid leukaemia (AML)<sup>29,30</sup>, chronic myeloid leukaemia (CML)<sup>31</sup>, breast cancer<sup>32</sup>, glioblastoma<sup>6,33</sup>, colorectal cancer<sup>34–36</sup>, pancreatic cancer<sup>37</sup> and ovarian cancer<sup>38–40</sup>. In these studies, the cells that formed tumours were rare. Nonetheless, in principle, tumorigenic cells could be common in some cancers that are hierarchically organized consistent with the cancer stem-cell model<sup>41</sup>.

Tumorigenic potential can only be tested in a permissive environment. A persistent concern is that there could be cancer cells that have the potential to contribute to tumour growth and disease progression in patients but do not have the opportunity to exhibit this potential in certain transplantation assays. A transplantation assay might underestimate the frequency of cancer cells with tumorigenic potential for many reasons. Human cells must be transplanted into highly immunocompromised mice to escape the powerful xenogeneic immune response that kills human cells in mice. Although a succession of increasingly immunocompromised mice have been used to assay the tumorigenic potential of human cancer cells, from nude to severe combined immunodeficient (SCID) to non-obese diabetic SCID (NOD/SCID) to NOD/SCID IL-2R $\gamma$ -null mice (NSG), all of these mice retain some xenogeneic immune barrier<sup>42</sup>. There is no opportunity to test whether human cells have tumorigenic potential if they are killed by a xenogeneic immune response. The frequencies of human AML cells<sup>30,43</sup>, acute lymphoblastic leukaemia (ALL) cells<sup>44,45</sup>, melanoma cells<sup>46</sup> and lung cancer cells<sup>47</sup> with leukaemogenic or tumorigenic activity are much higher in more highly immunocompromised mice.

Other mechanisms also contribute to an underestimation of tumorigenic potential. Many mouse malignant peripheral nerve sheath tumour (MPNST) cells have tumorigenic potential, but cells of one genotype depend on exogenous molecules of the protein laminin to form tumours, whereas cells of another genotype express laminin cell-autonomously<sup>48</sup>. This raises the arresting possibility that tumours from different patients might require different assay conditions to determine their full spectrum of tumorigenic cells. In other cases, key adhesion molecules or growth factors are required for cells to exhibit clonogenic

<sup>1</sup>Howard Hughes Medical Institute, Children's Research Institute and Department of Pediatrics, University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA.



**Figure 1 | Cancer cell fate versus potential.** **a**, Transplantation assays assess the potential of cancer cells to form tumours. The ability of a cell to form a tumour is context dependent: cells that can form a tumour under one set of conditions (environment one) may not form a tumour in other conditions (environment two). For this reason, tumorigenesis assays must be conducted under the most permissive possible conditions so as not to underestimate the spectrum of cells with tumorigenic potential. **b**, Lineage tracing or fate-mapping assays assess the actual fate of tumour cells in a

particular context, often the native tumour environment. Thus, whereas potential measures what a cell can do under permissive conditions, fate measures what a cell actually does in a particular context. Some cells with tumorigenic potential do not actually contribute to tumour growth — for example because they are in a non-permissive environment or because they are eliminated by immune effector cells (not shown). An important question is whether many (left) or few (right) cells with tumorigenic potential contribute to tumour growth.

activity but are not available to human cells in mouse tissues because of the inability of mouse ligands to bind human receptors<sup>49</sup>. Xenotransplanted tumours can also lack the architecture and stroma of tumours growing in native sites<sup>50</sup>. Many variables could influence the permissiveness of the environment for tumorigenesis, including uncharacterized factors that have not yet been taken into account in assays.

Given these concerns about xenotransplantation, it is reassuring that syngeneic transplantation of mouse cancer cells has also supported the cancer stem-cell model. Studies of cells from mouse germ-cell lineage cancers<sup>51</sup>, AML<sup>52,53</sup>, chronic myelomonocytic leukaemia (CMML)<sup>8</sup>, CML<sup>54</sup>, breast cancer<sup>55,56</sup> and medulloblastoma<sup>57,58</sup> have all been consistent with the cancer stem-cell model. Many of these studies demonstrated that some fractions of cancer cells are enriched for tumorigenic activity even when there is no xenogeneic barrier to engraftment. Therefore, the cancer stem-cell model cannot be entirely an artefact of xenotransplantation.

Improvements to transplantation assay conditions have revealed that some cancers have common leukaemogenic or tumorigenic cells (some AMLs<sup>59</sup>, many melanomas<sup>15,46,60</sup>, some ALLs<sup>59,61,62</sup> and mouse MPNSTs<sup>48</sup>), whereas other cancers continue to have only rare leukaemogenic or tumorigenic cells no matter what assay conditions are

used<sup>31,41,47</sup>. More work will be required to determine what fraction of cancers falls into each category.

For all of the reasons described, it will be crucial to continue to optimize transplantation assays to estimate as accurately as possible the spectrum of cancer cells that retain the potential to contribute to tumour growth. This will require systematically testing variations in xenotransplantation assay conditions, including the addition of variables such as human cytokines<sup>63</sup> that might influence human-cell engraftment<sup>64</sup>. Many researchers continue to transplant human cancers into heterotopic sites in mice without a careful comparison of the consequences of heterotopic compared with orthotopic engraftment. It will not be surprising if we are continuing to vastly underestimate the frequency of cells that can contribute to disease progression in some human cancers despite the advances made in xenotransplantation assays over the past few years.

### Tests for hierarchical organization

Beyond showing that a cancer has tumorigenic and non-tumorigenic cells, the other criterion that must be satisfied according to the cancer stem-cell model is that the tumorigenic cells give rise to non-tumorigenic progeny. Without demonstrating a lineage relationship,



the functional differences between cells may reflect genetic differences. Tumorigenic human cancer cells from AML<sup>29,30</sup>, CML<sup>31</sup>, breast cancer<sup>32</sup>, glioblastoma<sup>6,33</sup>, colorectal cancer<sup>34–36</sup>, pancreatic cancer<sup>37</sup> and ovarian cancer<sup>38–40</sup> formed more tumorigenic cells, as well as phenotypically distinct non-tumorigenic cells, in immunocompromised mice. For example, AML cells that carry the CD34 antigen (CD34<sup>+</sup>) but not the CD38 antigen (CD38<sup>−</sup>), from many patients, not only have enriched leukaemogenic activity but also form CD34<sup>−</sup> and CD38<sup>+</sup> non-leukaemogenic cells on transplantation<sup>65</sup>. Similar observations were made in mouse models of germ-cell lineage cancers<sup>51</sup>, AML<sup>52,53</sup>, CMML<sup>8</sup>, CML<sup>54</sup>, breast cancer<sup>55,56</sup> and medulloblastoma<sup>57</sup>. These findings are the basis for the idea that cancer stem cells form heterogeneous tumours by undergoing epigenetic changes, akin to the differentiation of normal stem cells.

However, for some cancers there is compelling evidence against the stem-cell model. Tumorigenic cells are common and phenotypically diverse in stage III and IV human melanomas<sup>15,46</sup>. One study<sup>66</sup> suggested that only CD271<sup>+</sup> melanoma cells can form tumours in immunocompromised mice. Another study<sup>67</sup> reported that CD271<sup>−</sup> melanoma cells can form tumours in NSG mice, but that they cannot form CD271<sup>+</sup> progeny or tumours in NOD/SCID mice. Our work suggests that both CD271<sup>−</sup> and CD271<sup>+</sup> melanoma cells readily form tumours in both NOD/SCID and NSG mice, and that CD271 is expressed heterogeneously in these tumours<sup>15,46</sup>. In an attempt to resolve the inconsistencies, we have compared the tumorigenic capacity of CD271<sup>+</sup> and CD271<sup>−</sup> melanoma cells isolated from multiple patients using our dissociation protocol<sup>15,46</sup>, as well as the dissociation protocols described by Civenni *et al.*<sup>67</sup> and Boiko *et al.*<sup>66</sup>. Irrespective of which dissociation protocol we used, whether or not we injected the cells with Matrigel, or whether we transplanted into NOD/SCID or NSG mice, both CD271<sup>+</sup> and CD271<sup>−</sup> melanoma cells readily formed tumours and all the tumours were heterogeneous for CD271 (S.J.M., unpublished observations). We were unable to detect any fraction of melanoma cells that lacks tumorigenic potential when we tested over 20 heterogeneously expressed markers in tumours obtained from many patients<sup>15,46</sup>.

Difficulty in reproducing cancer stem-cell markers has been a common problem in solid-cancer studies. For example, CD133 seemed to robustly distinguish tumorigenic from non-tumorigenic brain tumour cells in early studies<sup>6,33</sup>, but a series of subsequent studies<sup>68–71</sup> found tumorigenic cells in both CD133<sup>+</sup> and CD133<sup>−</sup> fractions. Because the existence of markers that can distinguish tumorigenic from non-tumorigenic cancer cells is the experimental basis for the conclusion that these cancers follow the stem-cell model, the inability to widely confirm many solid-cancer stem-cell markers undermines the evidence supporting the model.

A related problem is that cancer stem-cell markers that were originally characterized in a limited number of tumours have often been assumed to be generalizable. Such markers have frequently been used in other tumours, or even in cell lines, without independent confirmation that the markers were informative in these contexts. For example, it was clear from an initial study of breast cancer stem cells that the CD44<sup>+</sup>CD24<sup>−</sup> surface-marker phenotype enriched tumorigenic cells in some, but not all, breast cancers<sup>32</sup>. However, many studies subsequently characterized CD44<sup>+</sup>CD24<sup>−</sup> breast cancer stem cells in other tumours or cell lines without confirming in these contexts that the markers distinguished tumorigenic from non-tumorigenic cells. Consequently, various studies of cancer stem-cell properties are based on markers of uncertain validity.

### Variability among patients

One of the reasons for variability among studies is that markers expressed by tumorigenic cells differ among patients. Early studies on AML indicated that leukaemogenic cells were highly enriched among CD34<sup>+</sup>CD38<sup>−</sup> but not CD34<sup>+</sup>CD38<sup>+</sup> AML cells<sup>29,30</sup>. Studies with larger numbers of samples have shown that although the CD34<sup>+</sup>CD38<sup>−</sup> fraction consistently contains leukaemogenic cells, leukaemogenic activity is also commonly found in the CD34<sup>+</sup>CD38<sup>+</sup> and CD34<sup>−</sup> fractions<sup>43,65,72</sup>. About

half of AMLs have most leukaemogenic cells in the CD34<sup>+</sup>CD38<sup>−</sup> fraction and the other half have most leukaemogenic cells in the CD34<sup>+</sup>CD38<sup>+</sup> fraction<sup>43,65</sup>. Many human AMLs with nucleophosmin mutations have leukaemogenic activity exclusively in the CD34<sup>−</sup> fraction, but some have leukaemogenic activity in both CD34<sup>−</sup> and CD34<sup>+</sup> fractions<sup>73</sup>. This indicates that differences in mutations can cause differences in the phenotype of leukaemogenic cells among patients. The same is true for solid cancers. Sca1<sup>+</sup> cells have enriched tumorigenic activity in mouse lung adenocarcinomas with *Kras* and *p53* mutations but not in tumours with only *Kras* mutations<sup>74</sup>.

Differences among patients could also reflect differences in the cell of origin. For example, some medulloblastomas arise in the cerebellum from activation of the sonic hedgehog signalling pathway in granule neuron precursors and frequently have a poor prognosis<sup>75</sup>. Other medulloblastomas arise in the dorsal brainstem and are highly curable<sup>75</sup>. Similarly, neural progenitors from different regions of the central nervous system form different subtypes of ependymomas with different properties<sup>76</sup>. Both haematopoietic stem cells and restricted myeloid progenitors can serve as the cell of origin for AML<sup>53,77–79</sup>, but the leukaemogenic cells have somewhat different properties in each case<sup>80</sup>. The distinct developmental origins of tumours, both with respect to regional identity and position in the normal tissue hierarchy, contribute to differences among patients in tumorigenic cell properties.

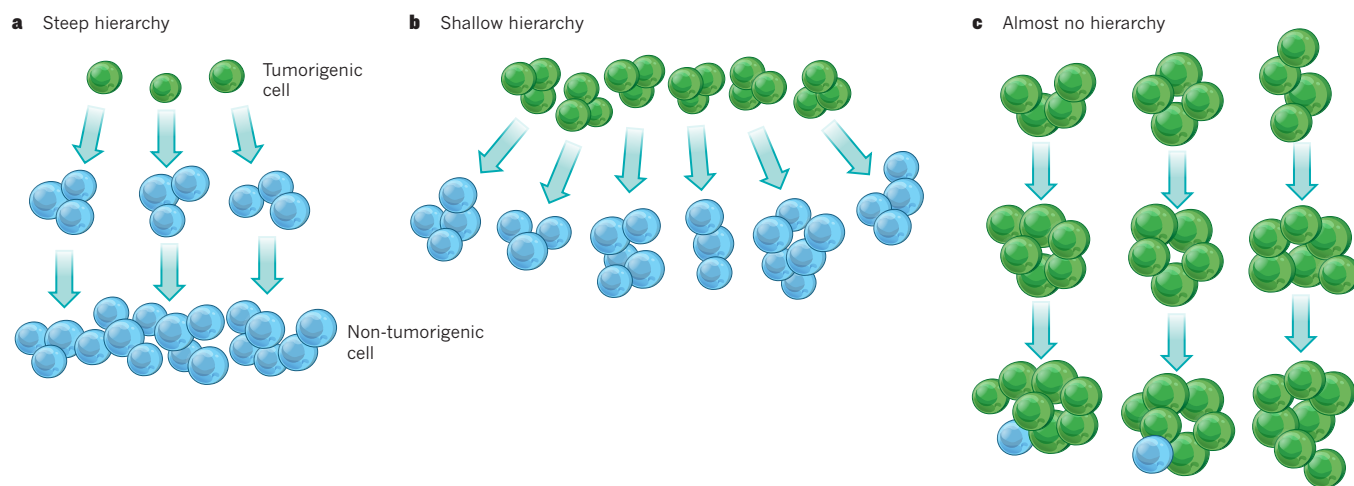
Tumorigenic cell phenotype can also change over time. In some ovarian cancers, only CD133<sup>+</sup> cells have tumorigenic activity, whereas in others tumorigenic cells are found in the CD133<sup>+</sup> and CD133<sup>−</sup> fractions<sup>40</sup>. Ovarian cancers with only CD133<sup>+</sup> tumorigenic cells sometimes give rise to CD133<sup>−</sup> tumorigenic cells on serial transplantation in mice<sup>40</sup>. If tumorigenic cell phenotypes commonly change on passaging of tumours, this could explain some of the inconsistencies observed among studies that use small numbers of tumours.

The frequency of tumorigenic cells in some cancers also varies widely among patients. Side-by-side studies of AMLs from different patients revealed frequencies of leukaemogenic cells in the CD34<sup>+</sup>CD38<sup>−</sup> cell fraction that varied by 1,000-fold<sup>65</sup>. B-lineage ALLs (B-ALLs) from different patients had frequencies of leukaemogenic cells that varied by 100-fold<sup>44</sup>. Ovarian cancers from different patients had tumorigenic cell frequencies that varied by almost 1,000-fold<sup>40</sup>. It remains uncertain to what extent this reflects biological variability in the frequency of cells that can contribute to tumour growth in patients as opposed to variability in the extent to which transplantation assays are permissive for tumorigenesis by cells of different genotypes. The variability in the frequency and identity of tumorigenic cells between patients shows that markers identified in one tumour cannot be assumed to distinguish cancer stem cells in other tumours or in other contexts.

A key question raised by the differences among patients is whether tumours of the same type differ in the extent to which they are hierarchically organized. For example, one possibility is that all breast cancers follow a stem-cell model even though existing markers do not distinguish tumorigenic from non-tumorigenic cells in some tumours. Another possibility is that only a subset of breast cancers follows the stem-cell model. Or that perhaps the hierarchy is steep in one subset of each cancer, with rare tumorigenic cells that give rise to abundant non-tumorigenic cells; and shallow in another subset of each cancer, with common tumorigenic cells that form a small number of non-tumorigenic cells (Fig. 2). Until enough tumours are carefully studied to observe consistent patterns among patients with particular subtypes of disease, tumours will have to be tested individually to determine whether tumorigenic cells are common or rare and whether markers can distinguish tumorigenic from non-tumorigenic cells.

### Fate versus potential in tumours *in vivo*

Which cells actually contribute to the growth and progression of tumours *in vivo*? Most cancer stem-cell studies are designed to assess the identity of cancer cells with the potential to contribute to tumour growth. But we also want to know the identity of cells fated to contribute to the growth and



**Figure 2 | Hierarchy of tumour cells.** Cancers may have a hierarchical organization in which subpopulations are arranged into tumorigenic and non-tumorigenic cells. **a**, Some hierarchies might be steep in which tumorigenic cells are rare but give rise to numerous non-tumorigenic cells. **b**, Other hierarchies might be shallow in which tumorigenic cells

are common but give rise to a small number of non-tumorigenic cells. **c**, Cancers may have almost no hierarchy, with very few non-tumorigenic cells. The shallower the hierarchy, the lower the value of distinguishing between tumorigenic and non-tumorigenic cells in order to understand cancer biology and improve therapy.

progression of specific tumours (Fig. 1). The question of fate addresses what cells actually do in a specific circumstance, whereas the question of potential addresses what cells can do under permissive conditions.

Some cells that have the potential to drive tumour growth do not actually do so in the native tumour because they are not in a permissive environment or they are killed by immune cells or by therapy. In the native tumour, slowly proliferating clones might be at a competitive disadvantage to rapidly proliferating clones and therefore may not contribute much to tumour growth. However, these slowly proliferating clones might form tumours after transplantation. Environmental cues from stromal cells can restrict the growth of cancer cells in the native tumour environment<sup>81</sup>, but the absence of these cues may permit the same cells to form tumours after transplantation. In some circumstances, a high percentage of cells with tumorigenic potential could contribute to tumour growth (Fig. 1). In other circumstances, there may only be a small percentage of cells with tumorigenic potential that actually contribute to tumour growth.

One context in which the fate of tumorigenic cells can be tracked is after xenotransplantation of human cancer cells. When primary human colorectal cancer cells were marked by lentiviral infection and the relative abundance of distinct clones was tracked as tumour cells were serially transplanted in mice<sup>82</sup>, individual clones differentially contributed to tumour growth over time. Some clones were always abundant, some were abundant then became rare and some were rare then became abundant. When human B-ALL cells from a single patient were transplanted at limiting and non-limiting cell doses, different dominant clones emerged in each recipient mouse<sup>44</sup>. Leukaemogenic clones thus do not contribute equally over time after transplantation.

Lineage tracing experiments in mouse models of benign tumours have tested whether many or few cells contribute to tumour growth<sup>20,21</sup>. In a recent study, a conditional reporter allele was used to track the fate of individual tumour cells in mice bearing benign papillomas. By inducing recombination of the conditional reporter allele in a small percentage of papilloma cells, rare marked clones were tracked over time to assess their contribution to tumour growth. The mice were treated with low doses of the oestrogen inhibitor tamoxifen, allowing keratin 14–CreER to permanently turn on a conditional reporter in a small percentage of papilloma cells<sup>20</sup>. The frequency of clones declined over time, with only 20% persisting after 7 weeks. Non-persisting clones seemed to be lost through terminal differentiation. The average number of cells in persisting clones increased over time: by 7 weeks the clones ranged from hundreds to thousands of cells. These observations demonstrate that

only a minor subpopulation of tumour cells drives papilloma growth, although the rate at which these cells divide is increased relative to that observed in the normal epidermis<sup>83</sup>. Benign papillomas are therefore hierarchically organized, consistent with the stem-cell model, although this is perhaps not surprising for a benign tumour.

The same study also fate-mapped cells in tumours after they progressed to squamous cell carcinomas<sup>20</sup>. Cells in these tumours were more highly proliferative, more undifferentiated and formed larger clones compared with cells in benign papillomas. A high percentage of cells contributed sustainably to the growth of squamous cell carcinomas, which, in this study, had only a shallow hierarchy, with few non-tumorigenic cells (Fig. 2). Additional studies are required to determine whether there is any hierarchical organization among the persisting clones — perhaps some have more proliferative potential than others. Overall, the data suggest that as benign adenomas progress to carcinomas, the hierarchy becomes shallower and more cells can contribute to tumour growth.

Lgr5 expression marks normal stem cells in the intestinal crypt<sup>84</sup>. To track the contribution of Lgr5<sup>+</sup> cells to the growth of premalignant intestinal adenomas, researchers have used a multicolour conditional reporter<sup>21</sup>. Administration of tamoxifen activated the Cre recombinase in Lgr5<sup>+</sup> cells, resulting in deletion of the APC tumour suppressor as well as activating the expression of a multicolour reporter in Lgr5<sup>+</sup> cells and their progeny<sup>21</sup>. Administration of a second pulse of tamoxifen reactivated the Cre recombinase, recombining the multicolour reporter again, leading to a colour switch in some Lgr5<sup>+</sup> cells in the adenomas. Lgr5<sup>+</sup> cells in the normal epithelium gave rise to adenomas and the Lgr5<sup>+</sup> cells within the adenomas contributed extensively to tumour growth. Most of the progeny were Lgr5<sup>−</sup> cells and so it was speculated that Lgr5<sup>+</sup> adenoma stem cells give rise to Lgr5<sup>−</sup> cells with little proliferative capacity; however, Lgr5<sup>−</sup> cells have not yet been fate-mapped, and it is unknown whether they have less ability than Lgr5<sup>+</sup> cells to contribute to tumour growth.

A recent study<sup>22</sup> shows that Lgr5<sup>−</sup> cells can also act as the cell of origin for intestinal adenomas in WNT-pathway activation and inflammation. Adenomas that arise in this context do not, on the basis of currently available data, seem to follow the cancer stem-cell model. The Lgr5<sup>−</sup> cells in these tumours give rise to Lgr5<sup>+</sup> cells, but both Lgr5<sup>−</sup> cells and Lgr5<sup>+</sup> cells can form clusters of cells called spheroids in culture and tumours *in vivo* with similar efficiency. This provides evidence for the argument that at least some intestinal adenomas are not hierarchically organized into Lgr5<sup>+</sup> tumorigenic cells and Lgr5<sup>−</sup> non-tumorigenic cells. Additional studies of



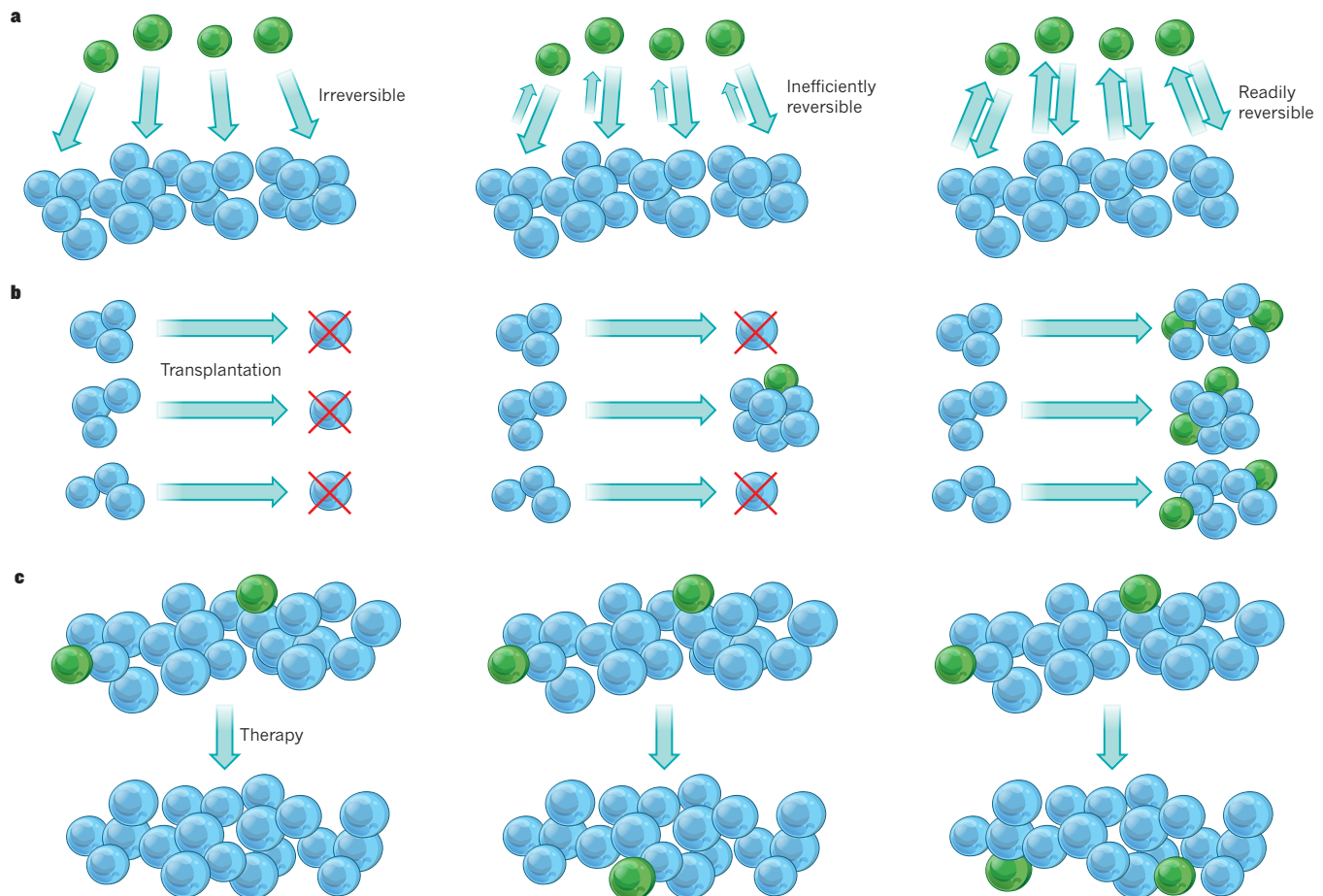
the Lgr5<sup>+</sup> and Lgr5<sup>-</sup> cell fractions from adenomas of other genetic backgrounds will be required to assess what fraction of adenomas exhibits hierarchical organization. It will also be important to test whether adenomas that do exhibit hierarchical organization continue to do so after they progress to malignancies.

These lineage tracing experiments thus provide limited support for the cancer stem-cell model. Although the growth of benign skin papillomas was driven by a minor subpopulation of cells, a much larger fraction of cells contributed to the growth of squamous cell carcinomas<sup>20</sup>. A key question now is whether some of the persistent clones that exhibit an ongoing contribution to tumour growth might nonetheless have limited tumorigenic potential in transplantation assays. The data on intestinal adenomas also offer limited support for the cancer stem-cell model as both Lgr5<sup>+</sup> and Lgr5<sup>-</sup> cells have the ability to serve as the cell of origin and to propagate tumours on transplantation, at least in certain genetic backgrounds. Ultimately, it will be necessary to integrate the data from both transplantation studies and fate-mapping studies of significant numbers of human and mouse tumours to understand the biological diversity.

### Fate testing through selective cell ablation

The selective ablation of genetically defined subsets of tumour cells is another approach to test which cells are fated to contribute to tumour

growth or disease progression in the native tumour environment. A recent study<sup>19</sup> addressed the role of Nestin<sup>+</sup> cells in the maintenance of a mouse model of glioma by ablating these cells. The protein GFP and the herpes simplex virus thymidine kinase (HSV-TK) were expressed under the control of the nestin promoter such that HSV-TK<sup>+</sup> cells could be selectively killed on administration of the antiviral ganciclovir. Nestin-expressing GFP<sup>+</sup> glioma cells were relatively quiescent and represented a minority of cells in the gliomas. Administration of ganciclovir modestly extended the lifespan of mice, indicating that the Nestin<sup>+</sup> cells contribute to tumour growth. Importantly, when tumours were reduced in size by treatment with the chemotherapeutic temozolomide, pulse-chase experiments suggested that tumour regrowth originated from the Nestin<sup>+</sup> fraction of tumour cells. Co-administration of temozolomide and ganciclovir significantly slowed tumour regrowth; however, it was impossible to assess the long-term effects of eliminating Nestin<sup>+</sup> cells because mice independently developed lethal tumours unrelated to the original tumour. Consequently, it remains unclear whether Nestin<sup>+</sup> cells are exclusively responsible for driving tumour growth and recurrence after therapy or whether Nestin<sup>-</sup> cells also contribute. It would be particularly interesting to selectively ablate Nestin<sup>-</sup> cells to determine whether this also slows tumour growth and extends mouse lifespan.



**Figure 3 | Predictions of transplantation and therapy response depend on the form of cell plasticity.** **a**, The differentiation of tumorigenic cells into non-tumorigenic progeny may be irreversible (left), inefficiently reversible (middle) or readily reversible (right). **b**, The degree of plasticity influences the outcome of transplantation assays. If differentiation is irreversible (left), non-tumorigenic cells should not form tumours after transplantation; if it is inefficiently reversible (middle), some tumours will form; and if it is readily reversible (right), tumours should form after transplantation. When cells efficiently and reversibly transition between tumorigenic and non-tumorigenic states, transplantation assays may not be able to distinguish

between cells in tumorigenic and non-tumorigenic states, and it might not be experimentally possible to test whether there is any hierarchical organization in the tumour. **c**, The predicted outcome of therapies designed to eliminate tumorigenic cells can also be influenced by plasticity in cancer cell hierarchies. If differentiation is irreversible (left), therapies that eliminate tumorigenic cells will convert a malignancy with hierarchical organization to a benign tumour containing only non-tumorigenic cells; if it is inefficiently reversible (middle), a single round of therapy will deplete but not eliminate tumorigenic cells; and if it is readily reversible (right), then a single round of therapy will have little effect on tumorigenic cell frequency.

## Reversible plasticity among cancer cells

Some cancer cells reversibly transition among states that differ in their competence to contribute to tumour growth<sup>85</sup>. For example, some cancer cells can reversibly transition between epithelial and mesenchymal states, and there is evidence that breast cancer cells in the mesenchymal state are more competent than those in the epithelial state to form tumours<sup>9</sup>. Melanoma cells reversibly turn on and off the histone demethylase JARID1B, and cells that express JARID1B are more competent to sustain tumour growth<sup>16</sup> than those that do not express JARID1B. Many other markers are reversibly turned on and off in lineages of melanoma cells in a manner that does not correlate with the ability to form a tumour<sup>15,17</sup>. Exposure of glioma cells to perivascular nitric oxide reversibly promotes their ability to form tumours<sup>13</sup>. The evidence that some cancer cells can undergo reversible changes in their competence to form tumours offers an alternative explanation for the increased tumorigenic potential of subsets of cancer cells that is independent of the differentiation of cancer stem cells.

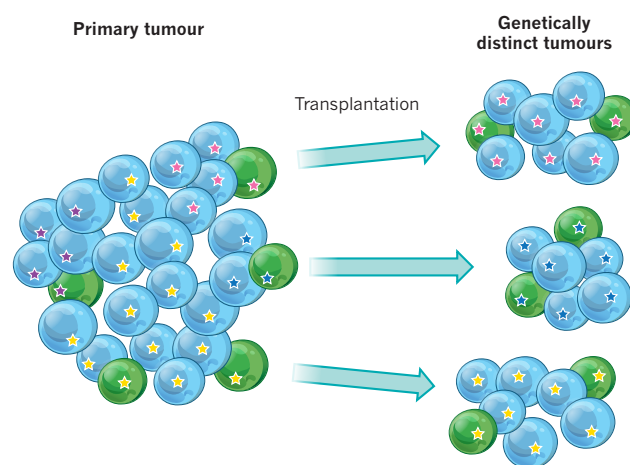
Drug resistance is also a plastic property of some cancer cells. Rare subpopulations in cancer cell lines that exhibit resistance to a variety of therapeutics reversibly form sensitive or resistant progeny depending on whether the cells are passaged with or without the therapeutic<sup>18</sup>. This raises the possibility of intrinsic therapy-resistance mechanisms that are not necessarily associated with a static hierarchy or an undifferentiated state.

It is crucial to distinguish models in which intrinsic differences in tumorigenic capacity reflect reversible changes in cell state from those in which intrinsic differences in tumorigenic capacity reflect irreversible differentiation, because these models make very different experimental and clinical predictions (Fig. 3). If the heterogeneity within tumours reflects cells that reversibly and efficiently transition between tumorigenic and non-tumorigenic states<sup>9,85</sup>, it may not be possible to experimentally identify any population that lacks tumorigenic potential (Fig. 3b). Furthermore, it would still be necessary to eliminate all cancer cells during therapy, as even the non-tumorigenic cells could drive disease recurrence by giving rise to tumorigenic cells (Fig. 3c). By contrast, if heterogeneity reflects hierarchical organization in which cancer stem cells irreversibly differentiate into non-tumorigenic cells, then therapies that eliminate cancer stem cells should be necessary and sufficient to cure disease (Fig. 3a, c). This distinction is thus crucial to understand the underlying biology and to develop more effective therapies. Importantly, almost all of the existing evidence for reversible transitions between tumorigenic and non-tumorigenic states comes from studies of cells in culture, often cell lines, so it remains uncertain to what extent reversible transitions occur between tumorigenic and non-tumorigenic states in spontaneously arising cancers *in vivo*.

## Genetic heterogeneity in tumours

The conclusion that cancer stem cells can recapitulate the heterogeneity of the tumours from which they are derived has consistently been based on analyses of small numbers of surface markers, calling into question the degree to which there is genetic heterogeneity within tumours that is not recapitulated after the transplantation of tumorigenic cells<sup>86</sup>. If genetic heterogeneity within tumours is low, then the differentiation of cancer stem cells into non-tumorigenic progeny could be the main driver of heterogeneity (Fig. 4). Conversely, if genetic heterogeneity is extensive, every tumorigenic cell could form a genetically distinct tumour rather than recapitulating the tumour from which it is derived. In tumours with extensive genetic heterogeneity, phenotypic and functional differences among cells cannot be assumed to reflect the differentiation of cancer stem cells — the variations could reflect genetic differences.

With these possibilities as a backdrop, it is interesting to consider the implications of recent deep-sequencing studies. Deep sequencing has been used to examine the genetic heterogeneity in tumours, the subclonal composition of tumours and the evolutionary relationships



**Figure 4 | Tumorigenic cells cannot recapitulate the heterogeneity of tumours with extensive genetic heterogeneity.** If every tumorigenic cell carries a combination of common and unique (coloured stars) mutations, then none of these cells will recapitulate the genetic heterogeneity of the tumour from which they are derived — they will all give rise to genetically distinct tumours on transplantation. They still may give rise to hierarchically organized tumours with tumorigenic and non-tumorigenic components, as in the tumour of origin. Nonetheless, if the genetic heterogeneity involves mutations that influence cancer cell phenotype or function, the genetic heterogeneity will contribute to tumour heterogeneity through mechanisms independent of cancer stem cell differentiation.

of mutations during disease progression. Deep sequencing cannot directly test the cancer stem-cell model. However, the frequencies of allelic variants in bulk tumour cells can be used to quantify the relative contribution of different clones to tumours. Even neutral, passenger, mutations can be informative because by following the contribution of the cells bearing these mutations to tumour growth and disease progression we gain insight into the fates of individual cancer cells and their progeny. Data from AML<sup>25,87</sup>, chronic lymphocytic leukaemia (CLL)<sup>88</sup>, breast cancer<sup>27,89,90</sup>, renal cell carcinoma<sup>28,91</sup> and pancreatic cancers<sup>92–94</sup> show surprisingly extensive genetic heterogeneity. Extensive genetic heterogeneity provides many opportunities for genetic changes to confer phenotypic and functional heterogeneity within tumours that is not addressed by the cancer stem-cell model (and that could complicate the testing of the model; Fig. 3).

It has long been known that cancer cells undergo clonal evolution in which mutations occur stochastically in individual cancer cells and are then subject to positive or negative selection depending on whether they confer a competitive advantage or disadvantage<sup>12,95</sup>. Cancer stem cells are no exception. Leukaemogenic ALL cells obtained from one patient exhibit genetic heterogeneity and undergo genetic changes over time when passaged in mice<sup>24,44</sup>. Whether these ALLs follow the cancer stem-cell model is not clear, because leukaemogenic cells are common in some ALLs and it has proven difficult to identify any clear hierarchical organization<sup>61,96</sup>. Human colorectal tumours have also been serially transplanted in mice, and their genetic heterogeneity assessed<sup>82</sup>. Only a small number of *de novo* genetic variants were detected in serially transplanted tumours compared with primary patient samples. Because there is compelling evidence that colorectal cancers are hierarchically organized into tumorigenic and non-tumorigenic components<sup>34–36</sup>, these results show that genetic changes do occur in colon cancer stem cells, although the paucity of such changes raises the possibility that the rate of mutagenesis might be suppressed in those cells. There is no inherent inconsistency between the cancer stem-cell model and the clonal evolution model<sup>4</sup> (Fig. 5).

## Therapy resistance

Tumorigenic cells in certain cancers are intrinsically resistant to certain therapies. For example, tumorigenic glioblastoma<sup>6</sup> and breast cancer<sup>7</sup> cells have been found to be enriched after irradiation of xenografts.



## BOX 1

# Challenges for evaluating the cancer stem–cell model

● Cancer stem-cell markers will have to be tested in significant numbers of patients<sup>15,40,65</sup> to account for heterogeneity among patients and to determine whether certain markers are more reliable in certain subsets of patients. This will provide insight into whether only certain subtypes or stages of disease follow the cancer stem-cell model. Studies should not assume that markers that distinguish tumorigenic from non-tumorigenic cells in one tumour will also do so in other tumours.

● To date, there is no evidence that any combination of cancer stem-cell markers isolates any cancer stem-cell population to a high degree of purity. The inability to purify any cancer stem cell is a profound impediment to characterizing the biology of these cells with precision. Until a high degree of purity can be demonstrated,

claims related to the cell-cycle distribution and gene expression profiles of ‘cancer stem cells’ will be of uncertain validity.

● It will be informative to perform lineage tracing<sup>20,21</sup> and selective cell ablation experiments<sup>19</sup> in other cancers and at other stages of disease to assess whether many, or few, cells are fated to contribute to tumour growth and disease progression.

● Studies of genetic heterogeneity should be integrated with studies of tumorigenic potential in order to develop a more realistic understanding of the extent to which individual tumorigenic cells recapitulate the heterogeneity of the tumour from which they are derived. In many cancers that follow the stem-cell model, tumorigenic cells may reproduce the cell hierarchy but not the genotype of the tumour of origin.

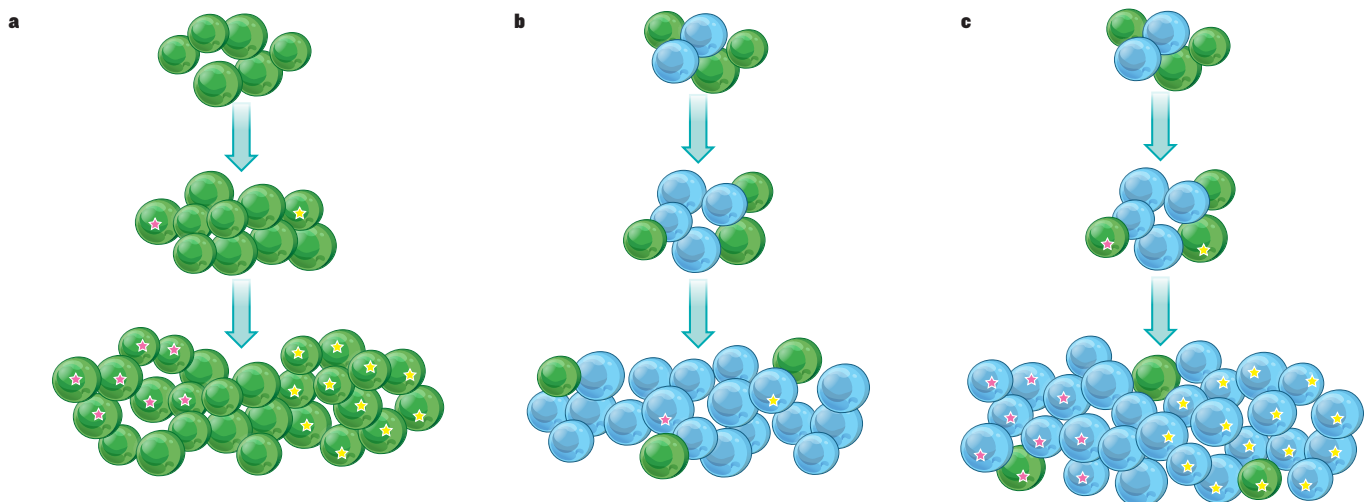
A similar enrichment of tumorigenic cells has been observed in cyclophosphamide-treated colorectal tumours<sup>97</sup>. Activation of ATM-dependent DNA-damage repair in tumorigenic glioblastoma cells<sup>6</sup> and decreased reactive oxygen species in tumorigenic breast cancer cells<sup>7</sup> may explain the therapy resistance of these cells.

Although it has been suggested that cancer stem cells can be defined by therapy resistance, this is not true in any general sense. Differentiation therapies specifically target cancer stem cells by exploiting their capacity to differentiate. Acute promyelocytic leukaemia (APL) is treated with arsenic trioxide and *trans*-retinoic acid to induce terminal differentiation, growth arrest and apoptosis by clonogenic APL cells<sup>98</sup>. Mouse glioblastoma stem cells can be induced to differentiate into glia by treatment with the protein BMP4, reducing proliferation, tumour growth and tumorigenic cell frequency<sup>99</sup>. BMP4 also promotes glial differentiation by normal central nervous system stem cells<sup>100</sup>, suggesting that tumorigenic cancer cells sometimes inherit differentiation pathways from normal stem cells in the same tissue. *Cis*-retinoic acid improves survival in high-risk neuroblastoma patients<sup>101,102</sup> by inducing the differentiation of

undifferentiated neuroblastoma cells. Thus, tumorigenic cells are more sensitive to some therapies and less sensitive to others compared with non-tumorigenic cells.

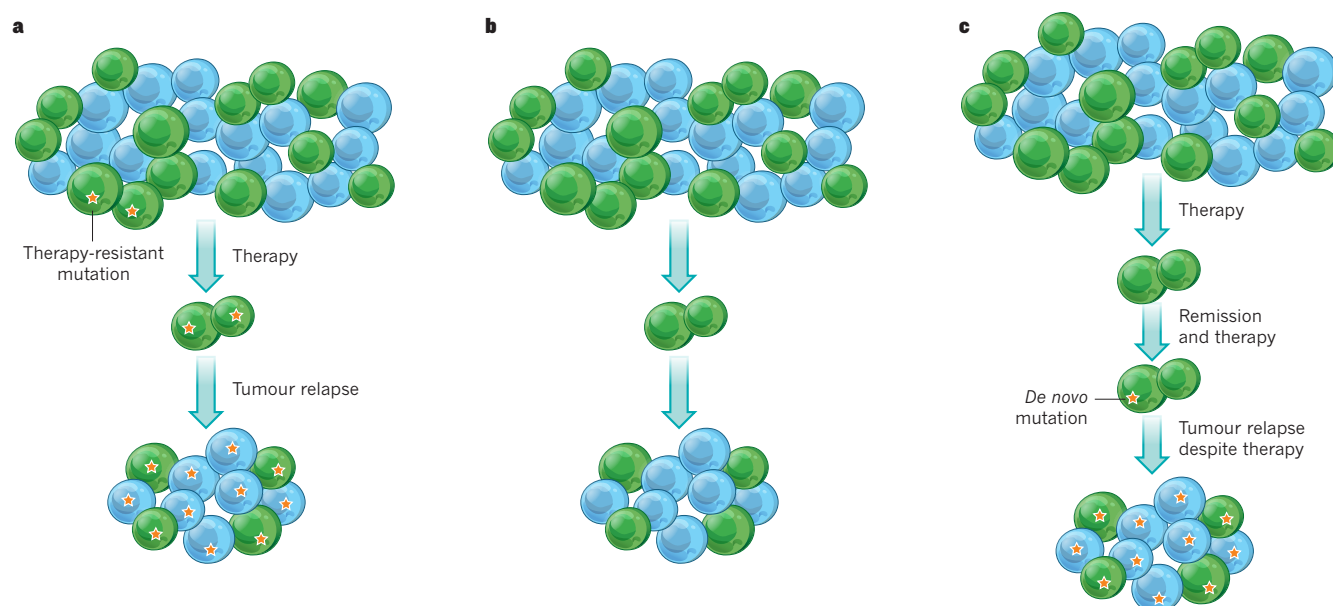
Genetic changes clearly confer therapy resistance in some circumstances<sup>103</sup>. Sequential genetic analysis of cancers before therapy and after relapse have been consistent with this. In ALL<sup>24,26</sup>, AML<sup>25</sup> and CLL<sup>88,104</sup>, minor subclones before therapy often become dominant after therapy. If the inherent therapy resistance of cancer stem cells were the main determinant of survival during therapy, then dominant clones before therapy would probably remain dominant after therapy. The observation that therapy selects for minor subclones suggests that survival is stochastic (many cells have a similarly low probability of surviving), non-cell-autonomously determined or determined by genetic differences among subclones.

When the dominant pretherapy clone remains dominant after relapse, the dominant clone often gains *de novo* mutations<sup>25,26,88</sup>. Relapse-specific mutations can confer therapy resistance. For example, recurrent relapse-specific mutations have been identified in the gene *NT5C2* in 10–20% of T-cell ALLs<sup>105,106</sup>. *NT5C2* is a 5' nucleotidase that



**Figure 5 | Clonal evolution and tumorigenic cell differentiation can independently or jointly contribute to tumour heterogeneity.** **a**, New mutations (different coloured stars) can increase the heterogeneity within tumours as long as the mutations influence cell phenotype or function. **b**, The differentiation of tumorigenic cells into non-tumorigenic progeny creates heterogeneity within tumours. New mutations that occur in non-tumorigenic cells would not be propagated (unless they restore tumorigenic potential).

**c**, If mutations occur in tumorigenic cells, then both clonal evolution and the differentiation of tumorigenic cells into non-tumorigenic progeny contribute to tumour heterogeneity. This is probably what occurs in cancers that follow the stem-cell model. This means that phenotypic and functional differences cannot automatically be ascribed to epigenetic differences among tumorigenic and non-tumorigenic cells as genetic heterogeneity may contribute to some of those differences.



**Figure 6 | Genetic changes and the inherent properties of tumorigenic cells can each contribute to therapy resistance.** **a**, Genetic alterations in some cells can confer therapy resistance (for example see ref. 105), whereby the altered cells persist after therapy and the tumour relapses. **b**, Tumorigenic cells in certain cancers are inherently resistant to certain therapies<sup>6,7,97</sup> and despite therapy killing most cells, some tumorigenic cells persist, and the tumour

relapses. **c**, Tumorigenic cells and genetic changes may both contribute to therapy resistance. Tumorigenic cells might persist despite therapy, but be unable to cause relapse because of an inability to regenerate significant numbers of non-tumorigenic cells in the presence of therapy. The acquisition of *de novo* mutations might enhance the cells' therapy resistance, enabling the tumour to relapse and the disease to progress.

can inactivate nucleoside analogues, such as the chemotherapy drugs 6-mercaptopurine and 6-thioguanine, which are used to treat ALL. The relapse-specific *NT5C2* mutations are predicted to be gain-of-function mutations that enhance enzymatic activity<sup>105,106</sup>. It is likely that certain genetic changes are sufficient for therapy resistance.

The intrinsic resistance of certain cancer stem cells to particular therapies combined with genetic change probably allows disease progression (Fig. 6). CML follows a cancer stem-cell model because CML stem cells form terminally differentiated myeloid cells during the chronic phase<sup>107</sup>. CML stem cells are inherently resistant to imatinib<sup>54,108–110</sup>. Cells with features of CML stem cells persist after therapy, even in patients who achieve a complete cytogenetic remission<sup>111</sup>. Imatinib is thought to restore the health of patients with CML mainly by eliminating differentiated CML cells. This profoundly reduces the number of leukaemia cells but leaves CML stem cells lurking in the patients, ready to re-establish the disease on discontinuation of imatinib<sup>112,113</sup>. As long as patients are maintained on imatinib they remain healthy, until a mutation arises in the fusion gene *BCR-ABL*, which confers imatinib resistance<sup>114,115</sup>. In this way, the inherent imatinib resistance of CML stem cells allows the disease to smoulder in remission, but a genetic change is required if resistance to imatinib is to become more robust and the disease is to progress in the face of therapy.

## Perspective

It has been suggested that the properties of cancer stem cells can explain diverse unsolved clinical problems, but in many cases these predictions have not been carefully tested (Box 1). The roles that other sources of heterogeneity (such as genetic) have in disease progression have, in many instances, not been factored into such suggestions. Moreover, difficulties verifying solid-cancer stem-cell markers have undermined efforts to confirm their existence in some cancers and to study their biology. It remains unclear to what extent these difficulties reflect variation in the expression of markers by cancer stem cells from different patients compared with the misguided efforts to apply the stem-cell model to cancers that lack hierarchical organization. The tendency not to publish data that are difficult to interpret or inconsistent with the model exaggerates the extent to which cancer stem-cell markers

are conserved among patients. It is time to critically test the model and its predictions, to acknowledge when the data do not fit the model and to integrate the data with other sources of heterogeneity when they do.

Tumorigenic cells are rare and phenotypically distinct in some cancers, but are common and phenotypically diverse with no clear hierarchical organization in others. We do not yet know what fraction of cancers follows the stem-cell model.

As we develop a more complete understanding of genetic heterogeneity within tumours, there may be some cancers in which genetic heterogeneity is the main driver of phenotypic and functional heterogeneity. In some cancers with pervasive genetic heterogeneity, it may not be possible to rigorously test the cancer stem-cell model as genetic differences within and between tumours could make it impossible to identify any reproducible hierarchical organization, even if there is differentiation into post-mitotic progeny. Some cancers may have epigenetic heterogeneity that is not well described by the cancer stem-cell model. Indeed, a general question concerns the extent to which the phenotypic and functional properties of cancer cells undergo reversible changes. New models of cancer heterogeneity and plasticity may emerge.

In cancers that do not follow the stem-cell model, it will be important to demonstrate this, to avoid fruitlessly focusing on small sub-populations of cancer cells that have no more capacity to drive disease progression or therapy resistance than other cancer cells. In cancers that do follow the stem-cell model, it will be important to clarify the markers that can be used to identify these cells and the contexts in which they work. It will also be crucial to integrate our understanding of the biology of these cells with our understanding of other sources of heterogeneity, to develop a realistic view of how each contributes to disease progression. For example, it is possible in some cancers that clones with a hierarchy of tumorigenic and non-tumorigenic cells may coexist in the same tumours with clones that have lost their hierarchical organization as a consequence of additional mutations. This would profoundly complicate the testing of the model and undermine the extent to which it can explain clinical behaviour. ■

Received 3 May; accepted 10 June 2013.



1. Dick, J. E. Stem cell concepts renew cancer research. *Blood* **112**, 4793–4807 (2008).
2. Kummermehr, J. & Trott, K.-R. in *Stem Cells* (ed. Potten, C. S.) 363–399 (Academic, 1997).
3. Reyes, T., Morrison, S. J., Clarke, M. F. & Weissman, I. L. Stem cells, cancer, and cancer stem cells. *Nature* **414**, 105–111 (2001).
4. Shackleton, M., Quintana, E., Fearon, E. R. & Morrison, S. J. Heterogeneity in cancer: cancer stem cells versus clonal evolution. *Cell* **138**, 822–829 (2009).
5. Clevers, H. The cancer stem cell: premises, promises and challenges. *Nature Med.* **17**, 313–319 (2011).
6. Bao, S. et al. Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. *Nature* **444**, 756–760 (2006).
7. Diehn, M. et al. Association of reactive oxygen species levels and radioresistance in cancer stem cells. *Nature* **458**, 780–783 (2009).
8. Oravecz-Wilson, K. I. et al. Persistence of leukemia-initiating cells in a conditional knockin model of an imatinib-responsive myeloproliferative disorder. *Cancer Cell* **16**, 137–148 (2009).
9. Mani, S. A. et al. The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* **133**, 704–715 (2008).
10. Balic, M. et al. Most early disseminated cancer cells detected in bone marrow of breast cancer patients have a putative breast cancer stem cell phenotype. *Clin. Cancer Res.* **12**, 5615–5621 (2006).
11. Magee, J. A., Piskounova, E. & Morrison, S. J. Cancer stem cells: impact, heterogeneity, and uncertainty. *Cancer Cell* **21**, 283–296 (2012).
12. Nowell, P. C. The clonal evolution of tumor cell populations. *Science* **194**, 23–28 (1976).
13. Charles, N. et al. Perivascular nitric oxide activates notch signaling and promotes stem-like character in PDGF-induced glioma cells. *Cell Stem Cell* **6**, 141–152 (2010).
14. Calabrese, C. et al. A perivascular niche for brain tumor stem cells. *Cancer Cell* **11**, 69–82 (2007).
15. Quintana, E. et al. Phenotypic heterogeneity among tumorigenic melanoma cells from patients that is reversible and not hierarchically organized. *Cancer Cell* **18**, 510–523 (2010).
- This article reports that many phenotypically diverse melanoma cells are capable of forming tumours that recapitulate the surface marker heterogeneity of the tumour from which they derive, suggesting that there may not be a hierarchy of tumorigenic and non-tumorigenic cells in melanoma.**
16. Roesch, A. et al. A temporarily distinct subpopulation of slow-cycling melanoma cells is required for continuous tumor growth. *Cell* **141**, 583–594 (2010).
17. Pinner, S. et al. Intravital imaging reveals transient changes in pigment production and Brn2 expression during metastatic melanoma dissemination. *Cancer Res.* **69**, 7969–7977 (2009).
18. Sharma, S. V. et al. A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. *Cell* **141**, 69–80 (2010).
- This article reports that the acquisition of tolerance to therapeutics is transient and reversible in lung cancer cell lines raising the possibility that there are epigenetic mechanisms of therapy resistance that do not necessarily have anything to do with stem-cell identity or static hierarchies.**
19. Chen, J. et al. A restricted cell population propagates glioblastoma growth after chemotherapy. *Nature* **488**, 522–526 (2012).
- Using a mouse model of malignant glioma, the authors of this study report that selective ablation of Nestin<sup>+</sup> glioma cells suggests that a relatively quiescent cell population sustains tumour growth after therapy.**
20. Driessens, G., Beck, B., Caauwe, A., Simons, B. D. & Blanpain, C. Defining the mode of tumour growth by clonal analysis. *Nature* **488**, 527–530 (2012).
- Lineage tracing in a mouse model of benign papilloma shows that only a fraction of papilloma cells sustainably contribute to tumour growth, although the frequency of such cells increases drastically on progression to squamous cell carcinoma.**
21. Schepers, A. G. et al. Lineage tracing reveals Lgr5<sup>+</sup> stem cell activity in mouse intestinal adenomas. *Science* **337**, 730–735 (2012).
- Fate mapping of Lgr5<sup>+</sup> cells in intestinal adenomas shows that these cells contribute to tumour growth while also forming Lgr5<sup>+</sup> progeny, raising the possibility that some adenomas are hierarchically organized.**
22. Schwitala, S. et al. Intestinal tumorigenesis initiated by dedifferentiation and acquisition of stem-cell-like properties. *Cell* **152**, 25–38 (2013).
- This study shows that both Lgr5<sup>+</sup> and Lgr5<sup>-</sup> cells within intestinal adenomas are capable of forming tumours that are heterogeneous for Lgr5, suggesting that some adenomas may not be hierarchically organized.**
23. Nakanishi, Y. et al. Dclk1 distinguishes between tumor and normal stem cells in the intestine. *Nature Genet.* **45**, 98–103 (2013).
24. Anderson, K. et al. Genetic variegation of clonal architecture and propagating cells in leukaemia. *Nature* **469**, 356–361 (2011).
25. Ding, L. et al. Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* **481**, 506–510 (2012).
- By sequencing samples from patients with AML before and after therapy, this study shows that new genetic variants commonly emerge after therapy, suggesting that therapy resistance is commonly determined by genetic variants.**
26. Mullighan, C. G. et al. Genomic analysis of the clonal origins of relapsed acute lymphoblastic leukemia. *Science* **322**, 1377–1380 (2008).
27. Navin, N. et al. Tumour evolution inferred by single-cell sequencing. *Nature* **472**, 90–94 (2011).
28. Xu, X. et al. Single-cell exome sequencing reveals single-nucleotide mutation characteristics of a kidney tumor. *Cell* **148**, 886–895 (2012).
29. Lapidot, T. et al. A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature* **367**, 645–648 (1994).
30. Bonnet, D. & Dick, J. E. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nature Med.* **3**, 730–737 (1997).
31. Wang, J. C. et al. High level engraftment of NOD/SCID mice by primitive normal and leukemic hematopoietic cells from patients with chronic myeloid leukemia in chronic phase. *Blood* **91**, 2406–2414 (1998).
32. Al-Hajj, M., Wicha, M. S., Benito-Hernandez, A., Morrison, S. J. & Clarke, M. F. Prospective identification of tumorigenic breast cancer cells. *Proc. Natl Acad. Sci. USA* **100**, 3983–3988 (2003).
33. Singh, S. K. et al. Identification of human brain tumour initiating cells. *Nature* **432**, 396–401 (2004).
34. Dalerba, P. et al. Phenotypic characterization of human colorectal cancer stem cells. *Proc. Natl Acad. Sci. USA* **104**, 10158–10163 (2007).
35. O'Brien, C. A., Pollett, A., Gallinger, S. & Dick, J. E. A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. *Nature* **445**, 106–110 (2007).
36. Ricci-Vitiani, L. et al. Identification and expansion of human colon-cancer-initiating cells. *Nature* **445**, 111–115 (2007).
37. Li, C. et al. Identification of pancreatic cancer stem cells. *Cancer Res.* **67**, 1030–1037 (2007).
38. Curley, M. D. et al. CD133 expression defines a tumor initiating cell population in primary human ovarian cancer. *Stem Cells* **27**, 2875–2883 (2009).
39. Zhang, S. et al. Identification and characterization of ovarian cancer-initiating cells from primary human tumors. *Cancer Res.* **68**, 4311–4320 (2008).
40. Stewart, J. M. et al. Phenotypic heterogeneity and instability of human ovarian tumor-initiating cells. *Proc. Natl Acad. Sci. USA* **108**, 6468–6473 (2011).
- This study demonstrates that CD133 enriches tumorigenic cells in only a subset of ovarian cancers and that even in these tumours the effectiveness of the marker changes on passaging.**
41. Kennedy, J. A., Barabe, F., Poepl, A. G., Wang, J. C. & Dick, J. E. Comment on “Tumor growth need not be driven by rare cancer stem cells”. *Science* **318**, 1722 (2007).
42. Shultz, L. D., Ishikawa, F. & Greiner, D. L. Humanized mice in translational biomedical research. *Nature Rev. Immunol.* **7**, 118–130 (2007).
43. Taussig, D. C. et al. Anti-CD38 antibody-mediated clearance of human repopulating cells masks the heterogeneity of leukemia-initiating cells. *Blood* **112**, 568–575 (2008).
44. Notta, F. et al. Evolution of human BCR-ABL1 lymphoblastic leukaemia-initiating cells. *Nature* **469**, 362–367 (2011).
45. Chiu, P. P., Jiang, H. & Dick, J. E. Leukemia-initiating cells in human T-lymphoblastic leukemia exhibit glucocorticoid resistance. *Blood* **116**, 5268–5279 (2010).
46. Quintana, E. et al. Efficient tumour formation by single human melanoma cells. *Nature* **456**, 593–598 (2008).
47. Ishizawa, K. et al. Tumor-initiating cells are rare in many human tumors. *Cell Stem Cell* **7**, 279–282 (2010).
48. Buchstaller, J., McKeever, P. E. & Morrison, S. J. Tumorigenic cells are common in mouse MPNSTs but their frequency depends upon tumor genotype and assay conditions. *Cancer Cell* **21**, 240–252 (2012).
49. Manz, M. G. Human-hemato-lymphoid-system mice: opportunities and challenges. *Immunity* **26**, 537–541 (2007).
50. Kuperwasser, C. et al. Reconstruction of functionally normal and malignant human breast tissues in mice. *Proc. Natl Acad. Sci. USA* **101**, 4966–4971 (2004).
51. Kleinsmith, L. J. & Pierce, G. B. Multipotentiality of single embryonic carcinoma cells. *Cancer Res.* **24**, 1544–1551 (1964).
52. Yilmaz, O. H. et al. Pten dependence distinguishes haematopoietic stem cells from leukaemia-initiating cells. *Nature* **441**, 475–482 (2006).
53. Krivtsov, A. V. et al. Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9. *Nature* **442**, 818–822 (2006).
54. Neering, S. J. et al. Leukemia stem cells in a genetically defined murine model of blast-crisis CML. *Blood* **110**, 2578–2585 (2007).
55. Vaillant, F. et al. The mammary progenitor marker CD61/β3 integrin identifies cancer stem cells in mouse models of mammary tumorigenesis. *Cancer Res.* **68**, 7711–7717 (2008).
56. Zhang, M. et al. Identification of tumor-initiating cells in a p53-null mouse model of breast cancer. *Cancer Res.* **68**, 4674–4682 (2008).
57. Read, T. A. et al. Identification of CD15 as a marker for tumor-propagating cells in a mouse model of medulloblastoma. *Cancer Cell* **15**, 135–147 (2009).
58. Ward, R. J. et al. Multipotent CD15<sup>+</sup> cancer stem cells in patched-1-deficient mouse medulloblastoma. *Cancer Res.* **69**, 4682–4690 (2009).
59. Kelly, P. N., Dakic, A., Adams, J. M., Nutt, S. L. & Strasser, A. Tumor growth need not be driven by rare cancer stem cells. *Science* **317**, 337 (2007).
60. Held, M. A. et al. Characterization of melanoma cells capable of propagating tumors from a single cell. *Cancer Res.* **70**, 388–397 (2010).
61. Williams, R. T., den Besten, W. & Sherr, C. J. Cytokine-dependent imatinib resistance in mouse BCR-ABL<sup>+</sup>, Arf-null lymphoblastic leukemia. *Genes Dev.* **21**, 2283–2287 (2007).
62. Rehe, K. et al. Acute B lymphoblastic leukaemia-propagating cells are present at high frequency in diverse lymphoblast populations. *EMBO Mol. Med.* **5**, 38–51 (2013).
63. Feuring-Buske, M. et al. Improved engraftment of human acute myeloid leukemia progenitor cells in β 2-microglobulin-deficient NOD/SCID mice and in NOD/SCID mice transgenic for human growth factors. *Leukemia* **17**, 760–763 (2003).

64. Notta, F., Doulatov, S. & Dick, J. E. Engraftment of human hematopoietic stem cells is more efficient in female NOD/SCID/IL-2Rgc-null recipients. *Blood* **115**, 3704–3707 (2010).
65. Eppert, K. *et al.* Stem cell gene expression programs influence clinical outcome in human leukemia. *Nature Med.* **17**, 1086–1093 (2011).
- The authors of this study demonstrated that the frequency and surface marker phenotype of leukaemic stem cells varied among samples from 16 patients.**
66. Boiko, A. D. *et al.* Human melanoma-initiating cells express neural crest nerve growth factor receptor CD271. *Nature* **466**, 133–137 (2010).
67. Civenni, G. *et al.* Human CD271-positive melanoma stem cells associated with metastasis establish tumor heterogeneity and long-term growth. *Cancer Res.* **71**, 3098–3109 (2011).
68. Chen, R. *et al.* A hierarchy of self-renewing tumor-initiating cell types in glioblastoma. *Cancer Cell* **17**, 362–375 (2010).
69. Joo, K. M. *et al.* Clinical and biological implications of CD133-positive and CD133-negative cells in glioblastomas. *Lab. Invest.* **88**, 808–815 (2008).
70. Wang, J. *et al.* CD133 negative glioma cells form tumors in nude rats and give rise to CD133 positive cells. *Int. J. Cancer* **122**, 761–768 (2008).
71. Beier, D. *et al.* CD133<sup>+</sup> and CD133<sup>−</sup> glioblastoma-derived cancer stem cells show differential growth characteristics and molecular profiles. *Cancer Res.* **67**, 4010–4015 (2007).
72. Sarry, J. E. *et al.* Human acute myelogenous leukemia stem cells are rare and heterogeneous when assayed in NOD/SCID/IL2Ryc-deficient mice. *J. Clin. Invest.* **121**, 384–395 (2011).
73. Taussig, D. C. *et al.* Leukemia-initiating cells from some acute myeloid leukemia patients with mutated nucleophosmin reside in the CD34<sup>+</sup> fraction. *Blood* **115**, 1976–1984 (2010).
74. Curtis, S. J. *et al.* Primary tumor genotype is an important determinant in identification of lung cancer propagating cells. *Cell Stem Cell* **7**, 127–133 (2010).
75. Gibson, P. *et al.* Subtypes of medulloblastoma have distinct developmental origins. *Nature* **468**, 1095–1099 (2010).
76. Johnson, R. A. *et al.* Cross-species genomics matches driver mutations and cell compartments to model ependymoma. *Nature* **466**, 632–636 (2010).
77. Cuzzio, A. *et al.* Similar MLL-associated leukemias arising from self-renewing stem cells and short-lived myeloid progenitors. *Genes Dev.* **17**, 3029–3035 (2003).
78. Somervaille, T. C. & Cleary, M. L. Identification and characterization of leukemia stem cells in murine MLL-AF9 acute myeloid leukemia. *Cancer Cell* **10**, 257–268 (2006).
79. Huntly, B. J. *et al.* MOZ-TIF2, but not BCR-ABL, confers properties of leukemic stem cells to committed murine hematopoietic progenitors. *Cancer Cell* **6**, 587–596 (2004).
80. Krivtsov, A. V. *et al.* Cell of origin determines clinically relevant subtypes of MLL-rearranged AML. *Leukemia* **27**, 852–860 (2013).
81. Hanahan, D. & Coussens, L. M. Accessories to the crime: functions of cells recruited to the tumor microenvironment. *Cancer Cell* **21**, 309–322 (2012).
82. Kreso, A. *et al.* Variable clonal repopulation dynamics influence chemotherapy response in colorectal cancer. *Science* **339**, 543–548 (2013).
83. Mascré, G. *et al.* Distinct contribution of stem and progenitor cells to epidermal maintenance. *Nature* **489**, 257–262 (2012).
84. Barker, N. *et al.* Identification of stem cells in small intestine and colon by marker gene *Lgr5*. *Nature* **449**, 1003–1007 (2007).
85. Gupta, P. B. *et al.* Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. *Cell* **146**, 633–644 (2011).
86. Marusyk, A., Almendro, V. & Polyak, K. Intra-tumour heterogeneity: a looking glass for cancer? *Nature Rev. Cancer* **12**, 323–334 (2012).
87. Walter, M. J. *et al.* Clonal architecture of secondary acute myeloid leukemia. *N. Engl. J. Med.* **366**, 1090–1098 (2012).
88. Schuh, A. *et al.* Monitoring chronic lymphocytic leukemia progression by whole genome sequencing reveals heterogeneous clonal evolution patterns. *Blood* **120**, 4191–4196 (2012).
89. Shah, S. P. *et al.* The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *Nature* **486**, 395–399 (2012).
90. Nik-Zainal, S. *et al.* The life history of 21 breast cancers. *Cell* **149**, 994–1007 (2012).
- Whole-genome sequencing of breast cancers revealed extensive genetic variation within tumours, suggesting that no tumorigenic breast cancer cell could recapitulate the heterogeneity of the tumour from which it derives.**
91. Gerlinger, M. *et al.* Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* **366**, 883–892 (2012).
92. Campbell, P. J. *et al.* The patterns and dynamics of genomic instability in metastatic pancreatic cancer. *Nature* **467**, 1109–1113 (2010).
93. Yachida, S. *et al.* Distant metastasis occurs late during the genetic evolution of pancreatic cancer. *Nature* **467**, 1114–1117 (2010).
94. Jones, S. *et al.* Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. *Science* **321**, 1801–1806 (2008).
95. Greaves, M. & Maley, C. C. Clonal evolution in cancer. *Nature* **481**, 306–313 (2012).
96. le Viseur, C. *et al.* In childhood acute lymphoblastic leukemia, blasts at different stages of immunophenotypic maturation have stem cell properties. *Cancer Cell* **14**, 47–58 (2008).
97. Dylla, S. J. *et al.* Colorectal cancer stem cells are enriched in xenogeneic tumors following chemotherapy. *PLoS ONE* **3**, e2428 (2008).
98. de Thé, H. & Chen, Z. Acute promyelocytic leukaemia: novel insights into the mechanisms of cure. *Nature Rev. Cancer* **10**, 775–783 (2010).
99. Piccirillo, S. G. *et al.* Bone morphogenetic proteins inhibit the tumorigenic potential of human brain tumour-initiating cells. *Nature* **444**, 761–765 (2006).
100. Gross, R. E. *et al.* Bone morphogenetic proteins promote astroglial lineage commitment by mammalian subventricular zone progenitor cells. *Neuron* **17**, 595–606 (1996).
101. Thiele, C. J., Reynolds, C. P. & Israel, M. A. Decreased expression of N-myc precedes retinoic acid-induced morphological differentiation of human neuroblastoma. *Nature* **313**, 404–406 (1985).
102. Matthay, K. K. *et al.* Treatment of high-risk neuroblastoma with intensive chemotherapy, radiotherapy, autologous bone marrow transplantation, and 13-*cis*-retinoic acid. *N. Engl. J. Med.* **341**, 1165–1173 (1999).
103. Garraway, L. A. & Janne, P. A. Circumventing cancer drug resistance in the era of personalized medicine. *Cancer Discov.* **2**, 214–226 (2012).
104. Landau, D. A. *et al.* Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell* **152**, 714–726 (2013).
105. Zzoneva, G. *et al.* Activating mutations in the *NT5C2* nucleotidase gene drive chemotherapy resistance in relapsed ALL. *Nature Med.* **19**, 368–371 (2013).
106. Meyer, J. A. *et al.* Relapse-specific mutations in *NT5C2* in childhood acute lymphoblastic leukemia. *Nature Genet.* **45**, 290–294 (2013).
107. Sawyers, C. L. Chronic myeloid leukemia. *N. Engl. J. Med.* **340**, 1330–1340 (1999).
108. Corbin, A. S. *et al.* Human chronic myeloid leukemia stem cells are insensitive to imatinib despite inhibition of BCR-ABL activity. *J. Clin. Invest.* **121**, 396–409 (2011).
109. Graham, S. M. *et al.* Primitive, quiescent, Philadelphia-positive stem cells from patients with chronic myeloid leukemia are insensitive to ST1571 *in vitro*. *Blood* **99**, 319–325 (2002).
110. Jiang, X. *et al.* Chronic myeloid leukemia stem cells possess multiple unique features of resistance to BCR-ABL targeted therapies. *Leukemia* **21**, 926–935 (2007).
111. Bhatia, R. *et al.* Persistence of malignant hematopoietic progenitors in chronic myelogenous leukemia patients in complete cytogenetic remission following imatinib mesylate treatment. *Blood* **101**, 4701–4707 (2003).
112. Rousselot, P. *et al.* Imatinib mesylate discontinuation in patients with chronic myelogenous leukemia in complete molecular remission for more than 2 years. *Blood* **109**, 58–60 (2007).
113. Savona, M. & Talpaz, M. Getting to the stem of chronic myeloid leukaemia. *Nature Rev. Cancer* **8**, 341–350 (2008).
114. Gorre, M. E. *et al.* Clinical resistance to STI-571 cancer therapy caused by BCR-ABL gene mutation or amplification. *Science* **293**, 876–880 (2001).
115. Shah, N. P. *et al.* Multiple BCR-ABL kinase domain mutations confer polyclonal resistance to the tyrosine kinase inhibitor imatinib (STI571) in chronic phase and blast crisis chronic myeloid leukemia. *Cancer Cell* **2**, 117–125 (2002).

**Acknowledgements** This work was supported by the Howard Hughes Medical Institute and the Cancer Prevention and Research Institute of Texas. We apologize to authors whose papers could not be cited owing to space restrictions.

**Author Information** Reprints and permission information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare competing financial interests: details accompany the full-text HTML version of this paper at [go.nature.com/yztadj](http://go.nature.com/yztadj). Readers are welcome to comment on the online version of this article at [go.nature.com/yztadj](http://go.nature.com/yztadj). Correspondence should be addressed to S. J. M. ([sean.morrison@utsouthwestern.edu](mailto:sean.morrison@utsouthwestern.edu)).



# The causes and consequences of genetic heterogeneity in cancer evolution

Rebecca A. Burrell<sup>1\*</sup>, Nicholas McGranahan<sup>1,2\*</sup>, Jiri Bartek<sup>3,4</sup> & Charles Swanton<sup>1,5</sup>

**Recent studies have revealed extensive genetic diversity both between and within tumours. This heterogeneity affects key cancer pathways, driving phenotypic variation, and poses a significant challenge to personalized cancer medicine. A major cause of genetic heterogeneity in cancer is genomic instability. This instability leads to an increased mutation rate and can shape the evolution of the cancer genome through a plethora of mechanisms. By understanding these mechanisms we can gain insight into the common pathways of tumour evolution that could support the development of future therapeutic strategies.**

Extensive genetic and phenotypic variation exist not only between tumours (intertumour heterogeneity) but also within individual tumours (intratumour heterogeneity; Fig. 1). In tumours, populations of genetically distinct subclones can intermingle or be spatially separated, and this subclonal architecture varies dynamically throughout the disease course<sup>1,2</sup>. Inter- and intratumour heterogeneity have significant implications for the choice of biomarkers to guide clinical decision-making in cancer medicine (see Siu and colleagues on page 355).

Genomic instability is a prominent source of genetic diversity within tumours, generating a diverse cell population that can be subject to selection in a given micro-environmental (see Review by Junttila and de Sauvage on page 346) or therapeutic context<sup>3–5</sup>. Genomic instability can arise through various routes, leaving distinct genomic footprints and differentially affecting tumour evolution and patient outcome. Different instability mechanisms can operate over the course of tumour development, and may be influenced by exposure to cancer drugs<sup>6,7</sup>. However, genomic instability might also be an attractive therapeutic target<sup>8</sup>.

In this Review we summarize the evidence for genetic and phenotypic variation both within and between tumours, and discuss how this might affect tumour biology, drug response and patient outcome. We focus on the role of genomic instability in generating inter- and intratumour heterogeneity, and review the array of instability mechanisms observed in cancer and their contribution to tumour evolution.

## Heterogeneity between different tumour types

It has long been appreciated that there are both genetic and functional differences between tumours. Tumours that originate from different tissues and cell types vary in terms of their genomic landscapes, prognosis and their response to cytotoxic therapies, probably resulting from the fact that the genetic events of transformation interact with cell-intrinsic biological properties. Reflecting this, the site of origin of the primary tumour frequently determines treatment decisions. However, considerable variation in terms of genomic aberrations, aggressiveness and drug sensitivity is also observed between tumours that originate from the same tissue and cell type<sup>9–11</sup>.

Mutational frequencies of oncogenes and tumour suppressors vary between tumours of different tissues, probably reflecting the importance of distinct signalling pathways within specific tissues or cellular contexts. Mutations in the tumour suppressor gene *APC*, for instance,

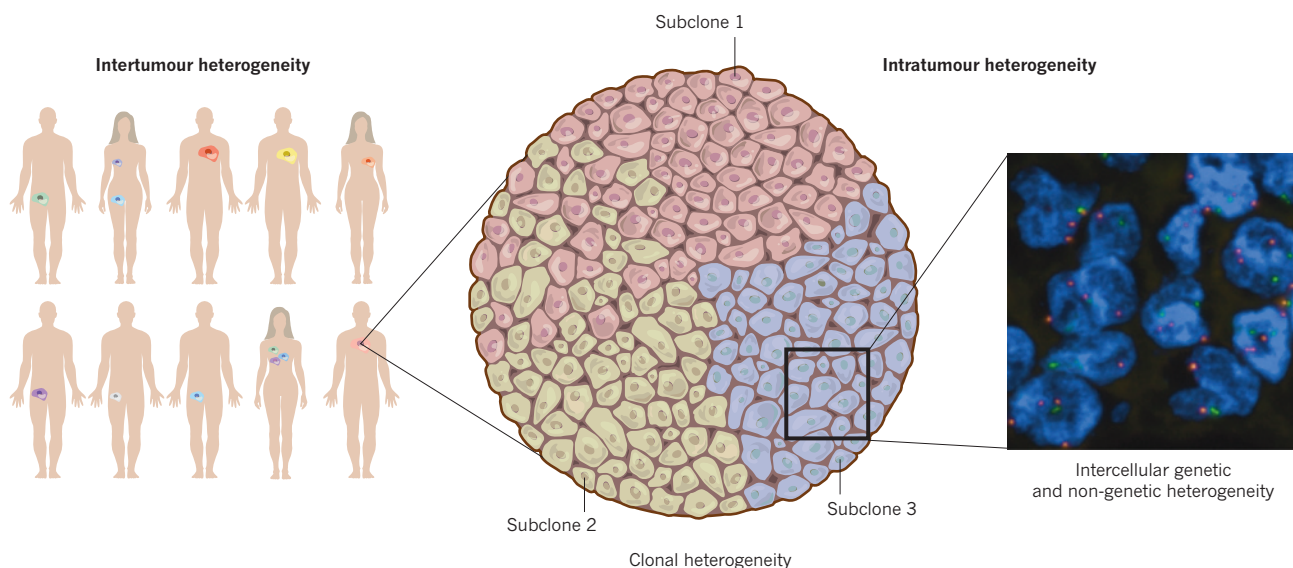
are common in gastrointestinal malignances such as colorectal cancer, probably owing to the importance of WNT-signalling pathways in intestinal epithelial homeostasis<sup>12</sup>. The specific cell type of origin can also profoundly influence tumour genetics and biology; for example, endocrine and exocrine pancreatic tumours exhibit considerable divergence in their mutational spectra and clinical behaviour<sup>13</sup>. Intrinsic differences between cell types may affect therapeutic response; the ability of cancer cells to undergo apoptosis in response to DNA-damaging agents seems to be substantially influenced by the cell type of origin<sup>14</sup>. Furthermore, drugs that target the mutant BRAF(V600E) oncoprotein seem to be less efficient in colorectal cancer cells than in melanoma cells<sup>15</sup>. This could be attributable to feedback upregulation of epidermal growth factor receptor (EGFR) after BRAF(V600E) inhibition in epithelial colorectal cancer cells but not in melanoma cells, which are derived from the neural crest and have lower basal EGFR expression<sup>15</sup>.

The clinical challenge that intertumour variation presents can be met to some extent by classifying tumours into subgroups. These may predict, for example, patient outcome or drug sensitivity on the basis of mutations, copy number changes, protein or RNA expression profiles, or patterns of genomic instability<sup>10,16,17</sup>. For example, breast cancers are routinely classified according to expression of hormone receptors as well as amplification of the gene *HER2* (also known as *ERBB2*) and expression of its protein<sup>10</sup>. Similarly, the discovery that around 10% of lung cancers harbour *EGFR*-activating mutations has led to the use of small molecule EGFR inhibitors in the treatment of this disease<sup>18</sup>.

However, recent advances in next-generation sequencing have revealed greater than expected inter- and intratumour genetic heterogeneity — very few mutations were observed in more than 5–10% of tumours of a particular tissue type<sup>9</sup>. In addition, within a cohort of tumours, the same gene may be affected by point mutation, DNA methylation, copy number alteration, or a combination of the three, underscoring the need for integrative approaches to analysing somatic aberrations in the cancer genome. Efforts to define tumour subgroups based on specific mutations may be confounded by epistasis, which involves the action of one gene on another. For example, in acute myeloid leukaemia (AML), *NPM1* mutations confer a favourable prognosis only in the presence of a co-occurring *IDH1* or *IDH2* mutation<sup>19</sup>.

Although genetic heterogeneity between tumours is known to be widespread, it has long been appreciated that phenotypic heterogeneity

<sup>1</sup>Translational Cancer Therapeutics Laboratory, Cancer Research UK London Research Institute, 44 Lincoln's Inn Fields, London WC2A 3LY, UK. <sup>2</sup>Centre for Mathematics and Physics in the Life Sciences and Experimental Biology (CoMPLEX), University College London, Physics Building, Gower Street, London WC1E 6BT, UK. <sup>3</sup>Danish Cancer Society Research Center, Strandboulevarden 49, DK-2100 Copenhagen, Denmark. <sup>4</sup>Institute of Molecular and Translational Medicine, Palacky University Olomouc CZ-775 15, Czech Republic. <sup>5</sup>UCL Cancer Institute, Paul O'Gorman Building, Huntley Street, London WC1E 6BT, UK. \*These authors contributed equally to this work.



**Figure 1 | Intertumour and intratumour heterogeneity.** Genetic and phenotypic variation are observed between tumours of different tissue and cell types, as well as between individuals with the same tumour type (intertumour heterogeneity). Within a tumour, subclonal diversity may be observed (intratumour heterogeneity). Subclones may intermingle (as shown by subclones 1 and 2) or be spatially separated (as shown by subclone 3). Separation between subclones could reflect physical barriers such as blood vessels or micro-environmental changes. Tumour subclones may show differential gene expression due to both genetic and epigenetic heterogeneity. Within a subclonal

population of tumour cells — shown here as a tumour section, hybridized to two fluorescent probes for the centromeres of two chromosomes (chromosome 2, red; chromosome 18, green) with DNA (blue) — there is intercellular genetic and non-genetic variation of, for example, chromosome copy number, somatic point mutations or epigenetic modifications that results in phenotypic diversity. Intercellular genetic heterogeneity is exacerbated by genomic instability, and may foster the emergence of tumour subclones. Genomic instability and tumour subclonal architecture may vary further over time if influenced by, for example, cancer treatment.

may not be as extensive, with the same cellular pathways often repeatedly affected<sup>9,20–24</sup>. Nevertheless, the impact of a given mutation will depend on the position and relative importance of the affected protein within the signalling pathway. An emerging theme from recent studies is that histone-modifying genes are recurrently mutated<sup>1,9,22,25</sup>, suggesting that altered chromatin regulation could be a common phenotype in a range of tumours. Another unifying feature of many genetically diverse malignancies is genomic instability — the same pattern of instability can occur through various distinct routes. For instance, deregulation of the G1–S transition, which is observed in almost all cancers, can be achieved through multiple means, simultaneously overriding cellular senescence, promoting proliferation and destabilizing the genome through DNA replication stress<sup>26–28</sup>.

Despite the evidence for functional convergence, it will nevertheless be challenging to identify biomarkers to define phenotypically similar, yet genetically diverse, lesions to guide treatment. In addition, the task of identifying and validating biomarkers may be complicated further by intratumour heterogeneity (See Review by Siu and colleagues on page 355).

### Intratumour heterogeneity and tumour evolution

Within tumours, diversity occurs in the expression of protein biomarkers<sup>10</sup>, and at multiple and different genetic<sup>29,30</sup> and epigenetic<sup>30,31</sup> levels. This diversity might complicate our ability to resolve the full spectrum of cancer pathway aberrations and confound biomarker validation through tumour sampling bias.

Genetically distinct subclonal populations of cells arise through intercellular genetic variation, followed by selective outgrowth of clones that have a phenotypic advantage within a given tumour micro-environmental context<sup>32–34</sup>. Fluctuations in subclonal architecture can occur, for example, in the context of disease progression (such as metastasis, see Review by Klein on page 365) or drug treatment<sup>1,6,21,35–37</sup>. A clonal sweep, whereby a new clone takes over the entire population, replacing ancestral clones, will result in a homogeneous cell population. During linear evolution, if a new clone fails to outcompete its predecessor, or

predecessors, a degree of heterogeneity will be observed<sup>38</sup>. Branched tumour evolution, in which distinct subclones evolve in parallel, however, will result in extensive subclonal diversity<sup>3</sup>.

A tumour's subclonal architecture can be reconstructed from sequencing approaches<sup>22,36,39–41</sup> or multiplex fluorescence *in situ* hybridization (FISH)<sup>2</sup>, providing insight into cancer evolution. Branched evolution has now been documented in a range of tumour types, including adenoma-to-carcinoma transition of the colon<sup>42</sup>, childhood acute lymphoblastic leukaemia (ALL)<sup>2</sup>, chronic lymphoblastic leukaemia (CLL)<sup>1</sup>, pancreatic cancer<sup>35</sup> and breast cancer<sup>39,41</sup>. In cases of ALL that exhibit branched evolutionary trajectories, heterogeneity has been observed between leukaemia-propagating cells<sup>2</sup>. This has important implications for cancer stem-cell models because it suggests that each subclonal branch was probably sustained by genetically distinct leukaemia-propagating cells<sup>2</sup>. It is conceivable that genetically heterogeneous subclonal lineages in solid tumours could also be maintained by distinct stem cells (See Review by Meacham and Morrison on page 328).

Branched evolutionary trajectories with spatial separation of tumour subclones (Fig. 1) have also been described. In clear-cell renal cell carcinoma, sequencing multiple biopsies from the same primary tumour revealed spatially separated subclones, harbouring heterogeneous somatic mutations and copy-number events<sup>40</sup>. Likewise, multiregion sampling in glioblastoma documented heterogeneous copy number events between different regions of the same tumour<sup>43</sup>. Evidence for clonal diversity between primary and metastatic sites has also been demonstrated in breast cancer<sup>36</sup>, pancreatic cancer<sup>21,35</sup> and medulloblastoma<sup>44</sup>. Intratumour heterogeneity of spatially separated subclones, at the DNA copy-number level, may be so profound that the copy number profiles of single tumour biopsies may more closely resemble those of tumours from different patients than those of adjacent biopsies of the same tumour<sup>45</sup>.

At metastatic sites, the new micro-environment, together with the selection pressures imposed by the process of metastasis itself, offer a plausible explanation for genetic differences between metastatic and primary sites<sup>46</sup> (see Perspective by Klein on page 365). However, it is



## BOX 1

# Main pathways implicated in genomic instability

● **Base and nucleotide excision repair.** These pathways function to excise and repair abnormal bases or nucleotides, such as deaminated cytosines or ultraviolet radiation induced pyrimidine dimers. Germline mutations in components of these pathways predispose people to colonic polyposis<sup>64</sup> or skin cancers<sup>66</sup>.

● **Mismatch repair.** This process acts during DNA replication to correct base mismatches, as well as insertions and deletions at repetitive sequences (microsatellites)<sup>68</sup>. Loss of function of *MSH2* and *MLH1*, which are required for mismatch repair, results in hypermutation and microsatellite instability. Mutations abrogating the proofreading ability of the DNA polymerases  $\delta$  and  $\epsilon$  increase base mismatches without affecting microsatellite stability<sup>12,67</sup>.

● **Telomere maintenance.** Telomere erosion or uncapping results in catastrophic chromosomal instability through end-on chromosome fusions, followed by bridging during mitosis and further breakage, in 'breakage–fusion–bridge' cycles. Reactivation of telomerase expression in cancer cells can alleviate this instability<sup>95</sup>.

● **Double-strand break repair.** Homologous recombination repair of double-strand breaks (DSBs) uses the sister DNA molecule as a template to repair the break, and is important for repair of stalled or collapsed replication forks. Non-homologous end joining directly ligates

the DSB in an error-prone manner. Defects in DSB detection or repair result in chromosomal instability<sup>8,81</sup>, with an elevated rate of structural chromosomal rearrangements<sup>124</sup> and point mutations<sup>70</sup>.

● **DNA replication.** Deregulated DNA replication can result in replication fork stalling, reversal and collapse. Deregulation can occur through oncogene activation, loss of certain tumour suppressors, DNA polymerase inhibition, nucleoside imbalances, replication-blocking DNA lesions and clashes of replication forks with ongoing transcription. Such replication stress can trigger DNA DSB formation (particularly at genomic fragile sites), unscheduled recombination events and chromosomal rearrangements<sup>26–28,87,97,125</sup>.

● **Chromosome segregation.** Defects in chromosome segregation can arise directly through defects in the mitotic checkpoint<sup>73,74</sup>, sister chromatid cohesion<sup>118</sup>, spindle geometry<sup>75,78,79</sup> and spindle dynamics<sup>76,77</sup>. These defects result in aberrant chromosome–spindle attachments, and missegregation of chromatids during anaphase, generating aneuploid daughter cells. Indirect segregation defects can be caused by structural chromosome rearrangements generated before mitosis, via replication stress<sup>87,89</sup>, defective repair<sup>90</sup> or telomere fusion<sup>80</sup>. Further damage to missegregating chromosomes can occur during cytokinesis<sup>85</sup> or aberrant replication and fragmentation of micronuclei<sup>86</sup>.

unclear precisely how spatial separation of genetically distinct clones arises in primary tumours<sup>40</sup>. It is possible that the separation reflects the presence of distinct micro-environmental niches in the primary tumour (see Junttila and de Sauvage on page 346), such that the sub-clones occupying each niche evolve relatively independently of one another. The similarity between this mode of tumour evolution and allopatric speciation (whereby geographically isolated populations from the same ancestral species evolve independently) raises the possibility that physical barriers in the tumour — such as blood vessels or tissue planes — prevent the mixing of tumour subclones (Fig. 1).

Nevertheless, intermingled heterogeneous clones are observed within single biopsies<sup>22,39,41,47</sup> to the extent that neighbouring cells can display amplification of different receptor tyrosine kinases — proteins that regulate key signalling processes and whose dysregulation has a crucial role in tumorigenesis<sup>48,49</sup>. At this intercellular level, genomic instability is thought to play a crucial part in generating enhanced genetic diversity. This intercellular heterogeneity increases phenotypic variation, broadening the pool of cells that are subject to selection, and therefore the likelihood of selective expansion of multiple different subclones and the emergence of a complex subclonal tumour architecture<sup>4</sup>.

## Genomic instability and tumour evolution

The genome is replicated and divided with high fidelity, such that the endogenous point-mutation rate in somatic cells is relatively low ( $0.77 \times 10^{-9}$  per site per cell division<sup>50</sup>) and errors in chromosome segregation are both rare (roughly 1 per 100 cell divisions) and poorly tolerated in non-transformed proliferating cells<sup>51–53</sup>. Disruption of mechanisms that maintain genome integrity or exposure to exogenous mutagens will elevate the mutation rate (mutation here is used in the broadest sense to include any genome aberration). Most solid tumours and haematopoietic malignancies display at least one form of genomic instability<sup>54</sup>.

Despite the clear association between genomic instability and cancer, there has been considerable debate as to whether elevated mutation rates are necessary for tumorigenesis<sup>55,56</sup>. Even with a normal somatic point mutation rate, the diversity and mutational burden within a large tumour cell population is likely to be extensive<sup>55</sup>, and sequencing

studies in AML and colorectal cancer have found that there is no need to invoke elevated mutation rates to explain the number of mutations observed<sup>20,38</sup>. Nevertheless, although not strictly necessary for carcinogenesis, the observation that most tumours display some form of instability suggests that it provides some benefit to tumour evolution. This hypothesis is supported by mouse models, which have revealed that genomic instability increases the risk of carcinogenesis<sup>57,58</sup>, leads to tumour regrowth after oncogene withdrawal<sup>59</sup> and promotes loss of heterozygosity of tumour suppressor genes<sup>60</sup>.

Different genomic instabilities result in different genetic lesions, ranging from increased point mutation frequency to small insertions and deletions, large-scale chromosomal rearrangements and alterations in ploidy (see Box 1 and Table 1). However, the phenotypic penetrance of these genetic changes may vary; a chromosome gain or loss is more likely to have functional consequences than a point mutation, most of which are neutral. As different patterns of genomic instability have distinct genomic footprints, it is possible to interrogate sequencing and copy-number data to examine how genomic instability has shaped tumour growth and evolution<sup>61–63</sup>. Given that most studies rely on examining a pooled population of cells, the true level of intratumour genetic heterogeneity driven by genomic instability is likely to be much greater than estimated, owing to both sampling bias and the inability to resolve heterogeneity between single cells. A recent study analysed an extensive catalogue of mutations in 30 of the most common cancer types and uncovered more than 20 signatures of processes that mutate DNA, driven by both exogenous and endogenous factors<sup>63</sup>.

In addition to global effects on genome architecture and mutational patterns, specific mutations may occur at an increased rate in the context of a particular instability mechanism. This is illustrated in colorectal tumours, in which those with defective mismatch repair exhibit a strikingly different mutational profile from those tumours with intact repair<sup>12</sup>. Similarly, colorectal tumours with defective base excision repair, due to *MYH* (also known as *MUTYH*) mutations (Box 1) show a preponderance of C-G→A-T transversions in *APC*<sup>64</sup>. In lung cancer, smoking increases the mutation rate and the frequency of C-G→A-T transversions<sup>22</sup> such that mutations in *TP53* can often be directly attributed to DNA damage from cigarette smoke carcinogens<sup>65</sup>. This effect

is likely to be compounded by epistasis, whereby a given mutation that occurs more frequently in the context of one pattern of instability may then lead to selection for specific combinations of mutations.

### Mechanisms of genomic instability

Increased point mutation frequencies can be caused by both exogenous factors<sup>22,63</sup> and defects in DNA repair pathways, such as mismatch repair and base and nucleotide excision repair<sup>12,64,66–68</sup> (see Box 1 and Table 1). In addition to hypermutation, mismatch repair deficiency is associated with expansion and contraction of repetitive tracts of DNA (microsatellite instability), leading to an increased frequency of frameshift mutations<sup>68</sup>. Elevated point mutation frequencies may also arise through mechanisms that affect gross chromosome structure, such as DNA replication stress<sup>69</sup> and defects in homologous recombination<sup>70</sup>.

Different mechanisms of instability can lead to particular distributions of point mutations — replication stress is associated with an enrichment of mutations in large genes<sup>69</sup>, whereas microsatellite instability may affect specific genes more frequently than others owing to the nucleotide composition of those genes<sup>68</sup>. Highly localized clusters of mutations that are associated with chromosomal rearrangement, known as kataegis, may be caused by the action of the APOBEC family of cytidine deaminases in breast cancer<sup>61</sup>. A recent study has highlighted the need to consider variation in mutation rate across the genome when identifying putative driver mutations; mutation frequency across the genome was found to strongly correlate with DNA replication timing and transcriptional activity<sup>11</sup>.

Chromosomal instability occurs at high frequencies across many cancer types and is associated with aggressive disease, drug resistance and poor prognosis<sup>54,71</sup>. This form of instability refers to an increased rate of change in chromosome number or structure, and results in a range of karyotypic abnormalities, including whole-chromosome and segmental aneuploidies, as well as translocations, inversions and deletions (Table 1). A substantial proportion of chromosomally unstable tumours also show increased ploidy, leading to suggestions that genome doubling could act as a precursor to chromosomal instability<sup>72</sup>.

Chromosomal instability can arise through a range of both mitotic and pre-mitotic defects. Mitotic defects result in inaccurate partitioning of

whole chromosomes between daughter cells — mediated through incorrect attachment of chromosomes to the mitotic spindle. This may occur through inadequate detection and correction<sup>73,74</sup>, or through increased formation, of incorrect attachments as a result of altered microtubule dynamics or transiently abnormal spindle geometry<sup>75–79</sup>. Pre-mitotic defects that converge on defective repair or increased generation of DNA damage can affect chromosome structural integrity. DNA damage may be acquired from exogenous sources (such as radiation), but in general chromosomal rearrangements in cancer cells are thought to arise through endogenous DNA damage. This may be, for example, through telomere dysfunction<sup>80</sup> or defects in DNA repair pathways (such as *BRCA1* and *BRCA2* mutations, which result in defective homologous recombination<sup>8,81</sup>; Box 1). A likely source of much of the DNA damage in tumours is DNA replication stress<sup>26,27</sup> (Box 1), which is reflected by the enrichment of chromosome breakpoints and loss of heterozygosity in cancer genomes at fragile sites (regions of the genome that are particularly susceptible to replication stress)<sup>82–84</sup>.

In recent years, there has been a growing appreciation that, from a mechanistic perspective, numerical and structural instability cannot be considered in isolation<sup>80,85–87</sup>. Mitotic defects in chromosome attachment, typically thought to lead to whole-chromosome aneuploidies, have been shown to result in DNA damage and structural rearrangements, through entrapment of chromosomes in the cytokinesis furrow<sup>85</sup> or through aberrant replication and fragmentation of micronuclei<sup>86</sup>. Conversely, structural chromosome rearrangements can also compromise the accuracy of chromosome segregation during mitosis<sup>80,87–90</sup>. Therefore, the relative contributions of mitotic defects and pre-mitotically generated structural chromosome abnormalities to chromosomal instability in tumours need to be explored.

### Variation in genomic instability over time

Tumour evolution is a dynamic process that contributes to both spatial and temporal tumour diversity. The type of genomic instability operating in a tumour at any point in time is likely to have a significant impact on tumour evolution, not only in terms of which specific pathways are dysfunctional<sup>91</sup>, but also in terms of the likelihood of a tumour

**Table 1 | Characteristic features of cancer genomes**

Genomic feature	Description	Possible mechanisms
Whole-genome duplication	Duplication of the entire DNA complement	Cell fusions, cytokinesis failure, metaphase or anaphase defects, or endoreduplication <sup>72</sup>
Chromosomal loss and gain	Loss and gain of whole chromosomes and segments of chromosomes	Mitotic defects: abnormal spindle geometry <sup>75,78,79</sup> , centriole amplification <sup>75,78</sup> , mitotic checkpoint defects <sup>73,74</sup> or hyperstable microtubules <sup>76,77</sup> Pre-mitotic defects: cohesin mutations <sup>118</sup> , DNA replication stress <sup>27,28,82,87</sup> , DNA damage repair defects <sup>81</sup> or telomere dysfunction <sup>80,95</sup>
Chromothripsis	Extensive genomic rearrangements involving one or more chromosomes	Ionizing radiation, micronucleus fragmentation or replication defects <sup>86,94,119</sup>
Translocations	Chromosomal rearrangements resulting in fusion of two chromosome arms	Non-homologous end joining, aberrant lymphocyte DNA recombination mechanisms and cytokinesis-induced DNA damage <sup>84,85,120</sup>
Tandem duplications	Sequential copies of a genomic segment that often spans several genes	Unknown <sup>121,122</sup>
Loss of heterozygosity at fragile sites	Copy number imbalance or gene conversion at genomic loci that are particularly sensitive to replication stress	DNA replication stress <sup>27,82</sup>
Focal deletions	Deletion of less than 5 megabases	Chromosome breakage or DNA replication stress <sup>82</sup>
Focal amplification	Amplification less than 5 megabases	Double minute chromosomes generated through chromosome breakage <sup>123</sup>
Hypermutation	Elevated mutation rate, leading to abundance of somatic mutations	DNA mismatch repair defects <sup>68</sup> , <i>BRCA2</i> mutations <sup>70</sup> and proofreading defects <sup>67</sup>
Kataegis	Localized hypermutation	AID/APOBEC activity <sup>61</sup>
C-G>T-A transitions	Base changes	Ultraviolet radiation-damage <sup>63</sup> , alkylating agents <sup>7</sup> and spontaneous deamination of methylated cytosines <sup>50</sup>
C-G>A-T transversions	Base changes	Base excision repair defects <sup>64</sup> , tobacco-related DNA damage <sup>22</sup> and cytotoxic chemotherapy <sup>6</sup>



metastasizing<sup>92</sup> and its response to treatment<sup>7,93</sup>.

Recent studies have revealed evidence for changing patterns of instability during tumour progression<sup>31,39,40</sup>. For example, in a study of clear-cell renal carcinoma, we found that chromosomally unstable metastases were probably derived from a tetraploid subclone in the primary tumour, whereas the remainder of the sampled regions in the primary tumour were diploid<sup>40</sup>. In a study of breast cancer, researchers made use of variant allele frequencies and copy number events that were derived from whole-genome sequencing reads to dissect the temporal sequence of mutations<sup>39</sup>. The mutations that occurred early in cancer development — on the trunk of the tumour's evolutionary tree — were found to exhibit a mutation pattern that was largely similar to germline mutations<sup>50</sup>. However, the mutations that occurred later during tumour progression had significantly altered mutational patterns, involving decreased frequencies of C-G→T-A base transitions<sup>39</sup>. These studies suggest that temporally separated biopsies may give strikingly different pictures of a tumour's genomic landscape.

Sequencing studies have also allowed researchers to analyse how tumour evolution and genomic instability could be affected by therapy. The analysis of pairs of primary and relapsed tumour cells from eight patients with AML found relapse-specific mutations and an increase in base transversions after chemotherapy, highlighting how treatment itself can act as a transient initiator of exogenous instability<sup>6</sup>. A study of nine glioblastomas from patients who had undergone chemotherapy with the alkylating agent temozolomide found that two cases exhibited an elevated frequency of mutations with a pattern that was consistent with induction by alkylating agents<sup>7</sup>. Both of these hypermutator gliomas relapsed after chemotherapy, and mutations in the mismatch repair gene *MSH6* (Box 1) were detected in the relapsed but not in pretreatment samples. This suggests that treatment with alkylating agents can impose selection for loss of mismatch repair function, allowing propagation of therapy-induced mutations. Because chromosomal instability is associated with multidrug resistance<sup>93</sup> and poor prognosis<sup>71</sup>, this raises the possibility that this form of genome instability may also be selected for during treatment.

It is important to note that observations of high numbers of mutations or chromosomal rearrangements do not necessarily imply ongoing genomic instability, but may instead reflect a historical prolonged period of mutational acquisition, or alternatively the occurrence of transient mutational bursts<sup>38,94</sup>. The evolution of the cancer genome may proceed not only in a gradual manner through clonal selection of point mutations or chromosomal alterations, but may also be subject to punctuated phases characterized by drastic changes. It has been postulated that chromothripsis, in which tens to hundreds of chromosomal rearrangements occur across one or a few chromosomes during a single catastrophic event, represents a major evolutionary leap in terms of tumour development<sup>94</sup>. Similarly, telomere dysfunction catalysed by eroded chromosome ends may be transient, leading to catastrophic instability until telomerase expression is activated to stabilize chromosome ends<sup>95</sup> (Box 1). Nevertheless, although these punctuated events are likely to contribute to cancer-genome evolution, observations of an active DNA damage response, and variation in chromosome copy number in both early stage and advanced malignancies, suggest that ongoing instability occurs in a substantial proportion of tumours<sup>27,96,97</sup>.

Drastic genomic change may also result from genome doubling, which has been estimated to occur at frequencies as high as 60% in some cancer types<sup>98</sup>. Some studies have proposed that the process is an intermediate stage on the path to aneuploidy, precipitating genetic and functional intratumour heterogeneity through chromosomal instability (see ref. 72). Although genome doubling has been observed as an early event in cancer development<sup>99</sup>, recent sequencing studies have found evidence for it at later stages in disease progression<sup>39</sup>. Genome doubling has been observed after, rather than before, the onset of chromosomal instability, indicating that there may be additional evolutionary benefits to genome doubling beyond catalysing chromosome reassortment<sup>39,98</sup>. Furthermore, genome doublings are proposed to play a major part during the evolution of eukaryotes, creating evolutionary opportunities

by providing new genetic material for mutation and selection to act upon<sup>100</sup>. In tumour evolution, it is possible that both genome doublings and tandem duplications (Table 1) might facilitate subfunctionalization of duplicated genes.

### Intertumour variation in genomic instability

There is considerable heterogeneity in both the patterns and extent of genomic instability that are prevalent in different tumour types. Microsatellite instability is observed most frequently in colorectal tumours; however, it occurs only in a minority of these tumours<sup>16</sup>. This instability seems to preferentially occur in tumours of the proximal colon, suggesting an association with epigenetic or environmental factors<sup>68</sup>. By contrast, defective homologous recombination is observed more frequently in breast and ovarian cancers<sup>8</sup>. A given pattern of instability can be caused by multiple mechanisms; for example, chromosomal instability may be driven by different defects in different tumours<sup>52,75,87</sup>. Furthermore, based on cell-line studies, multiple defects can operate simultaneously in a given tumour cell<sup>52,75,78,87</sup>.

Different instability mechanisms are also associated with variations in clinical outcome, for example, colorectal tumours with microsatellite instability are associated with a better patient prognosis compared with chromosomally unstable tumours<sup>17</sup>. Differences in disease outcome could relate to the profound infiltration of immune cells observed in microsatellite unstable tumours compared with colorectal cancers that display chromosomal instability. This immune infiltrate may be driven by an increased repertoire of neoantigens in these tumours as a result of mismatch repair deficiency<sup>101</sup>. Alternatively, the hypermutator phenotype of microsatellite-instability tumours could place them at the limit of tolerable instability and treatment with mutagenic chemotherapy may push them over the threshold<sup>4</sup>. Consistent with this idea, *in vitro*, microsatellite-instability colorectal cancer cell lines are more sensitive than chromosomally unstable cell lines to a range of different targeted and cytotoxic agents<sup>93</sup>. Whatever the explanation, these observations suggest that not all forms of genetic heterogeneity are equal in terms of their impact on patient outcome.

In addition to variation in the pattern of instability between tumours, variation also occurs in the degree of instability between tumours. For example, in breast cancer, oestrogen receptor positive (ER<sup>+</sup>) tumours display significantly lower karyotypic complexity (a genomic footprint of chromosomal instability), compared to ER<sup>-</sup> tumours<sup>96</sup>. This variation in the degree of chromosomal instability is associated with different outcomes; we have found evidence in ER<sup>-</sup> breast cancer, squamous non-small-cell lung cancer, and serous ovarian and gastric cancer that tumours within the highest quartile of chromosomal instability have a better prognosis than those in intermediate quartiles of chromosomal instability<sup>96,102</sup>. These data suggest that the benefits of genomic instability must be balanced against its costs, although it is unclear whether the link with prognosis is independent of treatment or whether therapy itself can push the level of instability beyond the tolerable threshold.

An elevated mutation rate not only increases the supply of potentially beneficial mutations but also deleterious ones, and there may be an optimal level of instability, which if exceeded impedes tumour growth, a scenario analogous to mutational meltdown in bacteria or error catastrophe in viruses<sup>4</sup>. Examples of excessive instability resulting in cell lethality can be observed in the context of mitotic checkpoint dysfunction<sup>37,103</sup> or multipolar anaphases in cells with extra centrosomes<sup>75</sup>. Cells may adapt to limit catastrophic instability by clustering extra centrosomes to form a bipolar spindle and avoid lethal multipolar division<sup>75,78</sup>, and so efforts to prevent centrosome clustering for therapeutic purposes to exacerbate instability may be effective<sup>54</sup>. Similarly, excessive instability in the context of *BRCA1* and *BRCA2* mutations may be averted by downregulation of the p53-binding protein 53BP1, to rescue, in part, homologous recombination repair<sup>104</sup>. By understanding cellular processes and pathways such as these — which are required for tolerance or averting excessive levels of genomic instability

— therapeutic opportunities may be revealed. This is exemplified by the efficacy of poly(ADP-ribose) polymerase inhibitors in breast and ovarian cancers that bear *BRCA1* and *BRCA2* mutations<sup>8,105,106</sup>. Other possible means of targeting genomic instability include inhibiting replication checkpoints in cells undergoing replication stress<sup>107</sup> and perturbing pathways that are essential for the survival of aneuploid cells<sup>51,53,93</sup>. Further research is required to identify new strategies to target genomic instability in cancer, which may represent an attractive approach to treating heterogeneous tumours.

### Consequences of intratumour heterogeneity

Similar to that observed between different tumours, some degree of phenotypic convergence may also occur between genetically distinct tumour subclones<sup>40</sup>. However, intratumour genetic heterogeneity also results in phenotypic diversity, and this has been documented in a range of tumour types, affecting clinically relevant parameters such as gene-expression signatures that reflect prognosis, and response to therapeutic agents<sup>6,40,48,108</sup>. In renal cancer, distinct transcriptomic profiles that reflect both good and poor prognosis have been described within one primary tumour<sup>40</sup>, and sampling of 11 glioblastomas revealed that many patients displayed multiple transcriptomic subtypes within different biopsies from the same tumour<sup>43</sup>.

Increasing evidence suggests that efforts to forecast outcome require the identification of low-frequency genetically and functionally distinct subclones at diagnosis<sup>1,37</sup>. This is illustrated by a recent study in chronic lymphocytic leukaemia, which revealed the selection of subclones that harbour distinct driver events during therapy<sup>1</sup>. The eight patients that harboured a detectable subclonal driver in their pretreatment samples showed significantly shorter time to re-treatment or death compared with the four patients with no evidence of subclonal driver events. Similarly, in a case of multiple myeloma, a subclone that was barely detectable at diagnosis was shown to be dominant towards the end of the disease<sup>37</sup>. These data underscore the need for longitudinal tumour sampling over the disease course and throughout treatment because the subclone that influences outcome may not be detectable in a single biopsy, or indeed even in multiple biopsies, taken before treatment.

Subclonal driver mutations result directly in phenotypic variation, for example in alterations to signal transduction pathways. In clear-cell renal cell cancer, a heterogeneous *MTOR* mutation (resulting in constitutive mTOR kinase activity *in vitro*) was associated with differential mTOR signalling pathway activation after 6 weeks of mTOR inhibitor therapy<sup>40</sup>. In a heterogeneous glioblastoma multiforme cell population harbouring mosaic amplification of *EGFR* and *PDGFRA*, combined inhibition of both kinases was necessary to attenuate PI(3)K pathway activity *in vitro*<sup>48</sup>. Intermingling of tumour cells harbouring amplification of distinct receptor tyrosine kinases (cells harbouring amplification of both genes were rare) hints at a possible functional codependency between the two genetically distinct populations<sup>48,49</sup>, as has been demonstrated in models using the fruitfly *Drosophila melanogaster*<sup>109</sup>. Intriguingly, *EGFR* mutant tissue, cell lines and stem cells release interleukin 6 (IL-6) and leukaemia inhibitory factor, which contribute to enhanced growth of adjacent subclones through activation of wild-type *EGFR*<sup>108</sup>, providing further evidence of possible codependencies between minor and major subclones in tumours.

However, it is important to note that phenotypic heterogeneity is not only mediated through genetic diversity. Genetically homogeneous subclones can behave in functionally distinct ways after exposure to chemotherapy, mediated by varied patterns of proliferation, with dormant resting cells surviving cytotoxic exposure<sup>110</sup>. Phenotypic heterogeneity is not determined solely through genetic distinctions between subclones, but also through stochastic events in gene expression and protein stability, epigenetic divergence and micro-environmental fluctuations<sup>30,31,110,111</sup>. Indeed, a study examining the fate of daughter cells derived from the same mother cell after antimetabolic drug exposure found that the fates of the sister cells seemed to be independent of each other<sup>112</sup>.

### Intratumour heterogeneity and treatment strategies

Observations of intratumour heterogeneity pose a challenge to targeted therapies and raise important questions regarding future drug-development strategies, which may benefit from considering whether clonally dominant events may be more effective targets than subclonal lesions<sup>29</sup>. Founder events might include *EGFR* and *KRAS* mutations in lung cancer, VHL loss-of-function through mutation, methylation or homozygous copy number loss in clear-cell renal cell carcinoma<sup>40</sup>, and *KRAS* and *BRAF* mutations in colorectal carcinoma<sup>113</sup>. Of these examples, only *EGFR* and *BRAF* mutations are currently directly targetable.

In addition, targeting clonally dominant genetic events is complicated by observations that a founder event in one tumour can be a later event in others. For example, in lung cancer, activating *EGFR* mutations were found to display intratumour heterogeneity and spatial discordance in 6.3–28% of cases<sup>114</sup>, suggesting that in these tumours *EGFR* mutation was not a founder event. Even in tumours in which *EGFR* mutations are founder events, the presence of low-frequency subclones harbouring the T790M gatekeeper mutation, which confers resistance to *EGFR* tyrosine kinase inhibitors, have been shown to be associated with early disease progression in patients treated with *EGFR* inhibitors<sup>114</sup>. In patients with colorectal cancer who have undergone anti-*EGFR* antibody therapy, mutations in *KRAS* (which confer drug resistance) were detectable in circulating tumour DNA 5–6 months after therapy in 9 out of 24 tumours, 3 of which developed multiple distinct *KRAS* mutations. Mathematical modelling suggested that low-frequency subclones harbouring these *KRAS* mutations were present in the tumour before treatment<sup>115</sup>. The same challenges are likely to affect the efficacy of other targeted therapies.

### Future directions

As genomics technologies advance further and become more affordable, our understanding of how genomic instability shapes tumour evolution will be greatly enhanced. Both mathematical and experimental modelling approaches are also likely to inform our understanding of how heterogeneity affects tumour progression<sup>5,110,115,116</sup>. Longitudinal tumour sampling approaches will be essential to decipher the impact of tumour heterogeneity on cancer evolution, and developing minimally invasive methods to profile heterogeneous tumour genomes will play a major part in deciphering changes in the genomic architecture of tumours in real time, exemplified by circulating free tumour DNA analysis<sup>115</sup>. In addition, there is a need to develop clinically useful measures of heterogeneity<sup>117</sup>. The prevalence of different patterns of instability will be more robustly assessed in large patient cohorts, and new patterns will probably be identified as the wealth of genomic data is analysed<sup>63</sup>. It will be important to examine in greater depth the relationship between the rate of instability and the complexity of subclonal architectures, as well as whether certain forms of genomic instability are associated with a greater degree of intratumour heterogeneity. There is a crucial need to understand mechanisms driving genomic instability so that therapeutic approaches to limit cancer diversity, adaptation and drug resistance can be developed. ■

Received 19 April; accepted 13 July 2013.

- Landau, D. A. *et al.* Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell* **152**, 714–726 (2013).  
**This article examines how subclonal driver mutations influence outcome.**
- Anderson, K. *et al.* Genetic variegation of clonal architecture and propagating cells in leukaemia. *Nature* **469**, 356–361 (2011).  
**An article demonstrates branched cancer evolution and genetic heterogeneity between leukaemia-propagating cells in ALL.**
- Greaves, M. & Maley, C. C. Clonal evolution in cancer. *Nature* **481**, 306–313 (2012).
- Cahill, D. P., Kinzler, K. W., Vogelstein, B. & Lengauer, C. Genetic instability and Darwinian selection in tumours. *Trends Biochem. Sci.* **24**, M57–M60 (1999).  
**An early reference placing genomic instability within the framework of Darwinian evolution, and recognizing that there may be distinct genomic footprints for different instability mechanisms.**
- Merlo, L. M., Pepper, J. W., Reid, B. J. & Maley, C. C. Cancer as an evolutionary and ecological process. *Nature Rev. Cancer* **6**, 924–935 (2006).



6. Ding, L. *et al.* Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* **481**, 506–510 (2012).  
**This article demonstrates the impact of treatment upon both tumour evolution and mutational spectra.**
7. Hunter, C. *et al.* A hypermutation phenotype and somatic *MSH6* mutations in recurrent human malignant gliomas after alkylator chemotherapy. *Cancer Res.* **66**, 3987–3991 (2006).
8. Lord, C. J. & Ashworth, A. The DNA damage response and cancer therapy. *Nature* **481**, 287–294 (2012).
9. Vogelstein, B. *et al.* Cancer genome landscapes. *Science* **339**, 1546–1558 (2013).
10. Weigelt, B. & Reis-Filho, J. S. Histological and molecular types of breast cancer: is there a unifying taxonomy? *Nature Rev. Clin. Oncol.* **6**, 718–730 (2009).
11. Lawrence, M. S. *et al.* Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* **499**, 214–218 (2013).
12. The Cancer Genome Atlas Research Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**, 330–337 (2012).
13. Jiao, Y. *et al.* *DAXX/ATRAX*, *MEN1*, and mTOR pathway genes are frequently altered in pancreatic neuroendocrine tumors. *Science* **331**, 1199–1203 (2011).
14. Brown, J. M. & Attardi, L. D. The role of apoptosis in cancer development and treatment response. *Nature Rev. Cancer* **5**, 231–237 (2005).
15. Prahallad, A. *et al.* Unresponsiveness of colon cancer to BRAF(V600E) inhibition through feedback activation of EGFR. *Nature* **483**, 100–103 (2012).
16. Lengauer, C., Kinzler, K. W. & Vogelstein, B. Genetic instability in colorectal cancers. *Nature* **386**, 623–627 (1997).
17. Popat, S., Hubner, R. & Houlston, R. S. Systematic review of microsatellite instability and colorectal cancer prognosis. *J. Clin. Oncol.* **23**, 609–618 (2005).
18. Pao, W. & Chmielecki, J. Rational, biologically based treatment of EGFR-mutant non-small-cell lung cancer. *Nature Rev. Cancer* **10**, 760–774 (2010).
19. Patel, J. P. *et al.* Prognostic relevance of integrated genetic profiling in acute myeloid leukemia. *N. Engl. J. Med.* **366**, 1079–1089 (2012).
20. Jones, S. *et al.* Comparative lesion sequencing provides insights into tumor evolution. *Proc. Natl Acad. Sci. USA* **105**, 4283–4288 (2008).
21. Yachida, S. *et al.* Distant metastasis occurs late during the genetic evolution of pancreatic cancer. *Nature* **467**, 1114–1117 (2010).
22. Govindan, R. *et al.* Genomic landscape of non-small cell lung cancer in smokers and never-smokers. *Cell* **150**, 1121–1134 (2012).
23. Hanahan, D. & Weinberg, R. A. The hallmarks of cancer. *Cell* **100**, 57–70 (2000).
24. Strauss, M., Lukas, J. & Bartek, J. Unrestricted cell cycling and cancer. *Nature Med.* **1**, 1245–1246 (1995).
25. Dalgleish, G. L. *et al.* Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. *Nature* **463**, 360–363 (2010).
26. Halazonetis, T. D., Gorgoulis, V. G. & Bartek, J. An oncogene-induced DNA damage model for cancer development. *Science* **319**, 1352–1355 (2008).
27. Bartkova, J. *et al.* DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis. *Nature* **434**, 864–870 (2005).
28. Bartkova, J. *et al.* Oncogene-induced senescence is part of the tumorigenesis barrier imposed by DNA damage checkpoints. *Nature* **444**, 633–637 (2006).
29. Swanton, C. Intratumor heterogeneity: evolution through space and time. *Cancer Res.* **72**, 4875–4882 (2012).
30. Marusyk, A., Almendro, V. & Polyak, K. Intra-tumour heterogeneity: a looking glass for cancer? *Nature Rev. Cancer* **12**, 323–334 (2012).
31. Varley, K. E., Mutch, D. G., Edmonston, T. B., Goodfellow, P. J. & Mitra, R. D. Intra-tumour heterogeneity of *MLH1* promoter methylation revealed by deep single molecule bisulfite sequencing. *Nucleic Acids Res.* **37**, 4603–4612 (2009).
32. Gatenby, R. A., Gillies, R. J. & Brown, J. S. Of cancer and cave fish. *Nature Rev. Cancer* **11**, 237–238 (2011).
33. Nowell, P. C. The clonal evolution of tumor cell populations. *Science* **194**, 23–28 (1976).
34. Cairns, J. Mutation selection and the natural history of cancer. *Nature* **255**, 197–200 (1975).
35. Campbell, P. J. *et al.* The patterns and dynamics of genomic instability in metastatic pancreatic cancer. *Nature* **467**, 1109–1113 (2010).
36. Shah, S. P. *et al.* Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. *Nature* **461**, 809–813 (2009).  
**This article demonstrates clonal diversity between primary and metastatic sites using next-generation sequencing.**
37. Keats, J. J. *et al.* Clonal competition with alternating dominance in multiple myeloma. *Blood* **120**, 1067–1076 (2012).
38. Welch, J. S. *et al.* The origin and evolution of mutations in acute myeloid leukemia. *Cell* **150**, 264–278 (2012).
39. Nik-Zainal, S. *et al.* The life history of 21 breast cancers. *Cell* **149**, 994–1007 (2012).
40. Gerlinger, M. *et al.* Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* **366**, 883–892 (2012).
41. Shah, S. P. *et al.* The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *Nature* **486**, 395–399 (2012).
42. Thirlwell, C. *et al.* Clonality assessment and clonal ordering of individual neoplastic crypts shows polyclonality of colorectal adenomas. *Gastroenterology* **138**, 1441–1454 (2010).
43. Sottoriva, A. *et al.* Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. *Proc. Natl Acad. Sci. USA* **110**, 4009–4014 (2013).
44. Wu, X. *et al.* Clonal selection drives genetic divergence of metastatic medulloblastoma. *Nature* **482**, 529–533 (2012).
45. Martinez, P. *et al.* Parallel evolution of tumor subclones mimics diversity between tumors. *J. Pathol.* **230**, 356–364 (2013).
46. Yates, L. R. & Campbell, P. J. Evolution of the cancer genome. *Nature Rev. Genet.* **13**, 795–806 (2012).
47. Navin, N. *et al.* Tumour evolution inferred by single-cell sequencing. *Nature* **472**, 90–94 (2011).
48. Szerlip, N. J. *et al.* Intratumoral heterogeneity of receptor tyrosine kinases EGFR and PDGFRA amplification in glioblastoma defines subpopulations with distinct growth factor response. *Proc. Natl Acad. Sci. USA* **109**, 3041–3046 (2012).
49. Snuderl, M. *et al.* Mosaic amplification of multiple receptor tyrosine kinase genes in glioblastoma. *Cancer Cell* **20**, 810–817 (2011).  
**This article reports intermingled heterogeneous subclones with mutually exclusive amplification of targetable receptor tyrosine kinases.**
50. Lynch, M. Rate, molecular spectrum, and consequences of human mutation. *Proc. Natl Acad. Sci. USA* **107**, 961–968 (2010).
51. Thompson, S. L. & Compton, D. A. Proliferation of aneuploid human cells is limited by a p53-dependent mechanism. *J. Cell Biol.* **188**, 369–381 (2010).
52. Thompson, S. L. & Compton, D. A. Examining the link between chromosomal instability and aneuploidy in human cells. *J. Cell Biol.* **180**, 665–672 (2008).
53. Torres, E. M., Williams, B. R. & Amon, A. Aneuploidy: cells losing their balance. *Genetics* **179**, 737–746 (2008).
54. Gordon, D. J., Resio, B. & Pellman, D. Causes and consequences of aneuploidy in cancer. *Nature Rev. Genet.* **13**, 189–203 (2012).
55. Tomlinson, I. P., Novelli, M. R. & Bodmer, W. F. The mutation rate and cancer. *Proc. Natl Acad. Sci. USA* **93**, 14800–14803 (1996).
56. Loeb, L. A. Mutator phenotype in cancer: origin and consequences. *Semin. Cancer Biol.* **20**, 279–280 (2010).
57. Weaver, B. A., Silk, A. D., Montagna, C., Verdier-Pinard, P. & Cleveland, D. W. Aneuploidy acts both oncogenically and as a tumor suppressor. *Cancer Cell* **11**, 25–36 (2007).  
**This article describes a mouse model that illustrates the complex relationship between the level of chromosomal instability and tumorigenesis.**
58. Liu, X. *et al.* Somatic loss of BRCA1 and p53 in mice induces mammary tumors with features of human BRCA1-mutated basal-like breast cancer. *Proc. Natl Acad. Sci. USA* **104**, 12111–12116 (2007).
59. Sotillo, R., Schwartzman, J. M., Socci, N. D. & Benezra, R. Mad2-induced chromosome instability leads to lung tumour relapse after oncogene withdrawal. *Nature* **464**, 436–440 (2010).
60. Baker, D. J., Jin, F., Jegannathan, K. B. & van Deursen, J. M. Whole chromosome instability caused by Bub1 insufficiency drives tumorigenesis through tumor suppressor gene loss of heterozygosity. *Cancer Cell* **16**, 475–486 (2009).
61. Nik-Zainal, S. *et al.* Mutational processes molding the genomes of 21 breast cancers. *Cell* **149**, 979–993 (2012).  
**References 39 and 61 demonstrate how deep sequencing can be used to elucidate the evolutionary history of tumours, and develop tools to identify mutational signatures using whole-genome sequencing data; they also describe a novel highly localized mutational process, kataegis**
62. Stephens, P. J. *et al.* The landscape of cancer genes and mutational processes in breast cancer. *Nature* **486**, 400–404 (2012).
63. Alexandrov, L. B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013).
64. Sieber, O. M. *et al.* Multiple colorectal adenomas, classic adenomatous polyposis, and germ-line mutations in *MYH*. *N. Engl. J. Med.* **348**, 791–799 (2003).
65. Pfeifer, G. P. *et al.* Tobacco smoke carcinogens, DNA damage and p53 mutations in smoking-associated cancers. *Oncogene* **21**, 7435–7451 (2002).
66. Dumaz, N., Drougard, C., Sarasin, A. & Daya-Grosjean, L. Specific UV-induced mutation spectrum in the p53 gene of skin tumors from DNA-repair-deficient xeroderma pigmentosum patients. *Proc. Natl Acad. Sci. USA* **90**, 10529–10533 (1993).
67. Palles, C. *et al.* Germline mutations affecting the proofreading domains of POLE and POLD1 predispose to colorectal adenomas and carcinomas. *Nature Genet.* **45**, 136–144 (2013).
68. Shah, S. N., Hile, S. E. & Eckert, K. A. Defective mismatch repair, microsatellite mutation bias, and variability in clinical cancer phenotypes. *Cancer Res.* **70**, 431–435 (2010).
69. Nikolaev, S. I. *et al.* A single-nucleotide substitution mutator phenotype revealed by exome sequencing of human colon adenomas. *Cancer Res.* **72**, 6279–6289 (2012).
70. Yang, D. *et al.* Association of BRCA1 and BRCA2 mutations with survival, chemotherapy sensitivity, and gene mutator phenotype in patients with ovarian cancer. *J. Am. Med. Assoc.* **306**, 1557–1565 (2011).
71. McGranahan, N., Burrell, R. A., Endesfelder, D., Novelli, M. R. & Swanton, C. Cancer chromosomal instability: therapeutic and diagnostic challenges. *EMBO Rep.* **13**, 528–538 (2012).
72. Ganem, N. J., Storchova, Z. & Pellman, D. Tetraploidy, aneuploidy and cancer. *Curr. Opin. Genet. Dev.* **17**, 157–162 (2007).
73. Cahill, D. P. *et al.* Mutations of mitotic checkpoint genes in human cancers. *Nature* **392**, 300–303 (1998).
74. Sotillo, R. *et al.* Mad2 overexpression promotes aneuploidy and tumorigenesis in mice. *Cancer Cell* **11**, 9–23 (2007).
75. Ganem, N. J., Godinho, S. A. & Pellman, D. A mechanism linking extra centrosomes to chromosomal instability. *Nature* **460**, 278–282 (2009).
76. Bakhoum, S. F., Genovese, G. & Compton, D. A. Deviant kinetochore microtubule dynamics underlie chromosomal instability. *Curr. Biol.* **19**, 1937–1942 (2009).
77. Bakhoum, S. F., Thompson, S. L., Manning, A. L. & Compton, D. A. Genome stability is ensured by temporal control of kinetochore-microtubule dynamics. *Nature Cell Biol.* **11**, 27–35 (2008).

**This article demonstrates direct attenuation of chromosomal instability in cancer cells, through overexpression of microtubule depolymerases.**

78. Silkworth, W. T., Nardi, I. K., Scholl, L. M. & Cimini, D. Multipolar spindle pole coalescence is a major source of kinetochore mis-attachment and chromosome mis-segregation in cancer cells. *PLoS ONE* **4**, e6564 (2009).
  79. Kaseda, K., McAinsh, A. D. & Cross, R. A. Dual pathway spindle assembly increases both the speed and the fidelity of mitosis. *Biol. Open* **1**, 12–18 (2012).
  80. Pampalona, J., Soler, D., Genescà, A. & Tusell, L. Whole chromosome loss is promoted by telomere dysfunction in primary cells. *Genes Chromosomes Cancer* **49**, 368–378 (2010).
  81. Hoeijmakers, J. H. Genome maintenance mechanisms for preventing cancer. *Nature* **411**, 366–374 (2001).
  82. Dereli-Öz, A., Versini, G. & Halazonetis, T. D. Studies of genomic copy number changes in human cancers reveal signatures of DNA replication stress. *Mol. Oncol.* **5**, 308–314 (2011).
  83. Bignell, G. R. *et al.* Signatures of mutation and selection in the cancer genome. *Nature* **463**, 893–898 (2010).
  84. Barlow, J. H. *et al.* Identification of early replicating fragile sites that contribute to genome instability. *Cell* **152**, 620–632 (2013).
  85. Janssen, A., van der Burg, M., Szuhai, K., Kops, G. J. & Medema, R. H. Chromosome segregation errors as a cause of DNA damage and structural chromosome aberrations. *Science* **333**, 1895–1898 (2011).
  86. Crasta, K. *et al.* DNA breaks and chromosome pulverization from errors in mitosis. *Nature* **482**, 53–58 (2012).
- References 85 and 86 highlight different mechanisms that contribute to the complex interplay between structural and numerical chromosomal instability.**
87. Burrell, R. A. *et al.* Replication stress links structural and numerical cancer chromosomal instability. *Nature* **494**, 492–496 (2013).
  88. Gisselsson, D. Classification of chromosome segregation errors in cancer. *Chromosoma* **117**, 511–519 (2008).
  89. Kawabata, T. *et al.* Stalled fork rescue via dormant replication origins in unchallenged S phase promotes proper chromosome segregation and tumor suppression. *Mol. Cell* **41**, 543–553 (2011).
  90. Chan, K. L., North, P. S. & Hickson, I. D. BLM is required for faithful chromosome segregation and its localization defines a class of ultrafine anaphase bridges. *EMBO J.* **26**, 3397–3409 (2007).
- This study provides evidence that pre-mitotic defects can result in chromosome segregation errors.**
91. Markowitz, S. *et al.* Inactivation of the type II TGF- $\beta$  receptor in colon cancer cells with microsatellite instability. *Science* **268**, 1336–1338 (1995).
  92. Gutenberg, A. *et al.* High chromosomal instability in brain metastases of colorectal carcinoma. *Cancer Genet. Cytogenet.* **198**, 47–51 (2010).
  93. Lee, A. J. *et al.* Chromosomal instability confers intrinsic multidrug resistance. *Cancer Res.* **71**, 1858–1870 (2011).
  94. Stephens, P. J. *et al.* Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell* **144**, 27–40 (2011).
  95. Artandi, S. E. & DePinho, R. A. Telomeres and telomerase in cancer. *Carcinogenesis* **31**, 9–18 (2010).
  96. Roylance, R. *et al.* Relationship of extreme chromosomal instability with long-term survival in a retrospective analysis of primary breast cancer. *Cancer Epidemiol. Biomarkers Prev.* **20**, 2183–2194 (2011).
  97. Gorgoulis, V. G. *et al.* Activation of the DNA damage checkpoint and genomic instability in human precancerous lesions. *Nature* **434**, 907–913 (2005).
  98. Carter, S. L. *et al.* Absolute quantification of somatic DNA alterations in human cancer. *Nature Biotechnol.* **30**, 413–421 (2012).
  99. Galipeau, P. C. *et al.* 17p (p53) allelic losses, 4N (G2/tetraploid) populations, and progression to aneuploidy in Barrett's esophagus. *Proc. Natl Acad. Sci. USA* **93**, 7081–7084 (1996).
  100. Taylor, J. S. & Raes, J. Duplication and divergence: the evolution of new genes and old ideas. *Annu. Rev. Genet.* **38**, 615–643 (2004).
  101. Ishikawa, T. *et al.* Tumor-specific immunological recognition of frameshift-mutated peptides in colon cancer with microsatellite instability. *Cancer Res.* **63**, 5564–5572 (2003).
  102. Birkbak, N. J. *et al.* Paradoxical relationship between chromosomal instability and survival outcome in cancer. *Cancer Res.* **71**, 3447–3452 (2011).
  103. Kops, G. J., Foltz, D. R. & Cleveland, D. W. Lethality to human cancer cells through massive chromosome loss by inhibition of the mitotic checkpoint. *Proc. Natl Acad. Sci. USA* **101**, 8699–8704 (2004).
  104. Bouwman, P. *et al.* 53BP1 loss rescues BRCA1 deficiency and is associated with triple-negative and BRCA-mutated breast cancers. *Nature Struct. Mol. Biol.* **17**, 688–695 (2010).
  105. Fong, P. C. *et al.* Inhibition of poly(ADP-ribose) polymerase in tumors from BRCA mutation carriers. *N. Engl. J. Med.* **361**, 123–134 (2009).
  106. Bryant, H. E. *et al.* Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature* **434**, 913–917 (2005).
  107. Ma, C. X., Janetka, J. W. & Piwnicka-Worms, H. Death by releasing the breaks: CHK1 inhibitors as cancer therapeutics. *Trends Mol. Med.* **17**, 88–96 (2011).
  108. Inda, M. M. *et al.* Tumor heterogeneity is an active process maintained by a mutant EGFR-induced cytokine circuit in glioblastoma. *Genes Dev.* **24**, 1731–1745 (2010).
  109. Wu, M., Pastor-Pareja, J. C. & Xu, T. Interaction between Ras(V12) and scribbled clones induces tumour growth and invasion. *Nature* **463**, 545–548 (2010).
  110. Kreso, A. *et al.* Variable clonal repopulation dynamics influence chemotherapy response in colorectal cancer. *Science* **339**, 543–548 (2013).
  111. Meads, M. B., Gatenby, R. A. & Dalton, W. S. Environment-mediated drug resistance: a major contributor to minimal residual disease. *Nature Rev. Cancer* **9**, 665–674 (2009).
  112. Gascoigne, K. E. & Taylor, S. S. Cancer cells display profound intra- and interline variation following prolonged exposure to antimetabolic drugs. *Cancer Cell* **14**, 111–122 (2008).
  113. Vakiani, E. *et al.* Comparative genomic analysis of primary versus metastatic colorectal carcinomas. *J. Clin. Oncol.* **30**, 2956–2962 (2012).
  114. Chen, Z. Y. *et al.* EGFR mutation heterogeneity and the mixed response to EGFR tyrosine kinase inhibitors of lung adenocarcinomas. *Oncologist* **17**, 978–985 (2012).
  115. Diaz, L. A. *et al.* The molecular evolution of acquired resistance to targeted EGFR blockade in colorectal cancers. *Nature* **486**, 537–540 (2012).
  116. Ding, L. *et al.* Genome remodelling in a basal-like breast cancer metastasis and xenograft. *Nature* **464**, 999–1005 (2010).
  117. Merlo, L. M. *et al.* A comprehensive survey of clonal diversity measures in Barrett's esophagus as biomarkers of progression to esophageal adenocarcinoma. *Cancer Prev. Res. (Phila.)* **3**, 1388–1397 (2010).
- This study develops tools to quantify intratumour heterogeneity, and relates these indices to patient outcome.**
118. Solomon, D. A. *et al.* Mutational inactivation of STAG2 causes aneuploidy in human cancer. *Science* **333**, 1039–1043 (2011).
  119. Liu, P. *et al.* Chromosome catastrophes involve replication mechanisms generating complex genomic rearrangements. *Cell* **146**, 889–903 (2011).
  120. Wang, J. H. Mechanisms and impacts of chromosomal translocations in cancers. *Front. Med.* **6**, 263–274 (2012).
  121. McBride, D. J. *et al.* Tandem duplication of chromosomal segments is common in ovarian and breast cancer genomes. *J. Pathol.* **227**, 446–455 (2012).
  122. Ng, C. K. *et al.* The role of tandem duplicator phenotype in tumour evolution in high-grade serous ovarian cancer. *J. Pathol.* **226**, 703–712 (2012).
  123. Hahn, P. J. Molecular biology of double-minute chromosomes. *Bioessays* **15**, 477–484 (1993).
  124. Xia, F. *et al.* Deficiency of human BRCA2 leads to impaired homologous recombination but maintains normal nonhomologous end joining. *Proc. Natl Acad. Sci. USA* **98**, 8644–8649 (2001).
  125. Bermejo, R., Lai, M. S. & Foiani, M. Preventing replication stress to maintain genome stability: resolving conflicts between replication and transcription. *Mol. Cell* **45**, 710–718 (2012).

**Acknowledgements** We would like to apologize to those whose work we have not been able to cite owing to space limitations. C.S. is funded by Cancer Research UK, The Medical Research Council, EU FP7 (projects PREDICT and RESPONSIFY), Prostate Cancer Foundation and the Breast Cancer Research Foundation (BCRF). J.B. is funded by the Danish Cancer Society and the European Commission (projects DResponse and Biomedreg). We thank E. Grönroos and P. Gorman for the use of the tumour section image in Fig. 1.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [go.nature.com/8sjatd](http://go.nature.com/8sjatd). Correspondence should be addressed to C.S. ([charles.swanton@cancer.org.uk](mailto:charles.swanton@cancer.org.uk)).



# Influence of tumour micro-environment heterogeneity on therapeutic response

Melissa R. Junttila<sup>1</sup> & Frederic J. de Sauvage<sup>1</sup>

**Tumour formation involves the co-evolution of neoplastic cells together with extracellular matrix, tumour vasculature and immune cells. Successful outgrowth of tumours and eventual metastasis is not determined solely by genetic alterations in tumour cells, but also by the fitness advantage such mutations confer in a given environment. As fitness is context dependent, evaluating tumours as complete organs, and not simply as masses of transformed epithelial cells, becomes paramount. The dynamic tumour topography varies drastically even throughout the same lesion. Heterologous cell types within tumours can actively influence therapeutic response and shape resistance.**

In healthy tissue, the stroma functions as the main barrier against tumorigenesis; however, the presence of transformed tumour cells initiates crucial changes that can convert this environment into one that supports cancer progression. The orchestration of these changes involves recruitment of fibroblasts, migration of immune cells, matrix remodelling and eventually development of vascular networks. How does the genetic and phenotypic variation that exists within tumours, or intra-tumoral heterogeneity, influence tumour growth? The identification of genetic variability within the same tumour suggests complicated events of branched evolution. Regional differences in selective pressures such as hypoxia, acidity and the presence of growth factors exist within a tumour and actively shape its development. Conceivably, distinct environmental landscapes within a given tumour select for mutations that engender survival and expansion, thereby creating tumour cell heterogeneity.

The role of tumour stroma in carcinogenesis has been reviewed extensively elsewhere<sup>1,2</sup>. In this Review, we briefly define the tumour cell extrinsic compartments and discuss how they contribute to tumour heterogeneity. We then review therapeutic-resistance mechanisms that implicate stromal cell types within tumours. Last, we discuss clinical attempts to target the tumour micro-environment and the challenges that lie ahead in order to achieve durable responses when treating cancer patients.

## Origin and influence of micro-environment heterogeneity

Although the expansion of neoplastic cells generates the initial insult that instigates the creation of the tumour niche, non-transformed cell types within the milieu co-evolve with the tumour cells, so that both continuously participate in the process of tumorigenesis (Fig. 1).

## Cancer-associated fibroblasts

Fibroblasts are an abundant mesenchyme-derived cell type that maintain the structural framework in tissues. Quiescent fibroblasts differentially respond to damage, such as wounding, and become activated to support repair. Although normal fibroblasts typically suppress tumour formation<sup>3</sup>, cancer-associated fibroblasts (CAFs) can significantly promote tumorigenesis<sup>4</sup>. Compared with normal tissue fibroblasts, CAFs have increased proliferation, enhanced extracellular matrix production and unique cytokine secretion (for example, stromal cell-derived factor 1, SDF1; vascular endothelial growth factor, VEGF; platelet-derived growth factor, PDGF; and hepatocyte growth factor, HGF)<sup>5</sup>. Other mesenchyme-derived cell types, such as adipocytes, can also contribute to tumour growth and progression.

Phenotypic and functional heterogeneity occurs in healthy fibroblasts

and CAFs<sup>6,7</sup>. Differences in fibroblast behaviour and response lead to extensive tissue remodelling mediated by augmented expression of proteolytic enzymes (for example, matrix metalloproteinases), deposition of extracellular matrix and pathogenic angiogenesis by liberating pro-angiogenic factors within the matrix<sup>8</sup>. Heterogeneity may be attributable to unique damage signals to which fibroblasts are exposed or possibly to their origin<sup>9</sup>. Significant cell plasticity also exists within this cell population, as both mesenchymal-to-epithelial transitions, and epithelial-to-mesenchymal transitions are known to occur, further enhancing stromal heterogeneity.

Although extensive dissection of stromal fibroblast intra- and inter-tumoral heterogeneity (variation between tumours) is impaired by the lack of specific markers, CAFs within tumours are clinically relevant. For example, the abundance of stromal cells correlates with poor prognosis for several forms of cancer, including breast<sup>10</sup> and pancreatic<sup>11</sup> cancer. Elevated expression of matrix metalloproteinases correlates with increased aggressiveness and poor prognosis in certain cancers<sup>12</sup>. A significant link between increased adipose tissue and cancer risk has also been demonstrated, although the mechanisms are still being elucidated<sup>13</sup>.

## Vasculature

The tumour vascular network is dynamic and can limit tumour growth<sup>14</sup>. Vascular networks are derived through formation of new vessels (angiogenesis), co-option and modification of existing vessels within tissue, or recruitment and differentiation of endothelial precursors from bone marrow (vasculogenesis), all of which contribute to vascular heterogeneity in and among tumours. Vessel formation involves degradation and reincorporation of existing vascular basement membranes that vary in a tissue-specific manner<sup>15</sup>. In addition, because tissue-specific vascular function and signalling are documented in normal organ homeostasis<sup>16</sup>, this is probably also true in the tumour context. For example, it is conceivable that co-opted vessels maintain some tissue-specific characteristics, and the host tissue within which the tumour develops influences the resulting vascular network.

Uneven vascularization and differences in vascular maturity combined with a lack of drainage due to poor lymphatic vessel coverage contribute to the complex topography and variable interstitial pressure within tumours. Inadequate function of poorly organized tumour vasculatures results in areas of hypoxia and limited nutrient supply. Distance from vascular beds creates a gradient that has been shown to be crucial for the distribution of drugs to all cells in the tumour<sup>17</sup>. Such variations in the

<sup>1</sup>Department of Molecular Biology, Genentech, 1 DNA Way, South San Francisco, California 94080, USA.

vascular networks can generate distinct micro-environments within the tumour and contribute to inter- and intratumour heterogeneity, and ultimately influence clinical outcome. Microvessel density has been reported to be a significant prognostic factor for poor outcome in non-small-cell lung cancer (NSCLC)<sup>18</sup> and colorectal<sup>19</sup> and breast<sup>20</sup> cancer. Moreover, elevated expression of the predominant pro-angiogenic ligand VEGFA has been shown to be associated with a worse prognosis than those that have low expression of VEGFA in metastatic colorectal, lung and renal cell cancers<sup>21</sup>.

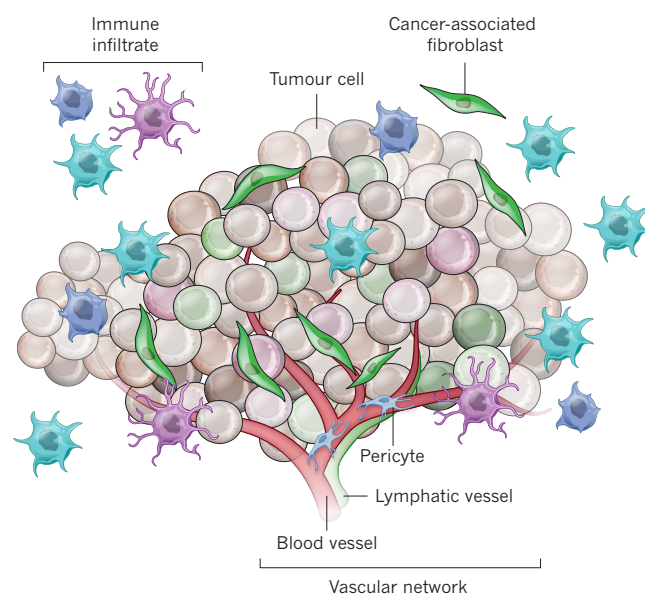
### Immune cells

The immune system collectively functions to recognize and protect tissues from infections and damage. Both the innate and adaptive immune systems have been implicated in promoting and preventing tumour growth. Although the immune system has the ability to mount anti-tumour responses, mechanisms of immune suppression can prevent this process.

Immune cell recruitment and localization in the tumour milieu vary widely in and among lesions. Heterogeneity of tumour immune contexture is influenced by various factors, including those secreted by CAF, the extent and permeability of the vasculature, and the tumour cells themselves. For instance, tumours with microsatellite instability, such as colorectal cancer subsets, may generate more neoantigens, leading to increased T-cell infiltration<sup>22</sup>. Similarly, the vascular bed within a tumour may strongly influence the immune contexture because endothelial cells regulate immune cell migration, as suggested by the unique endothelial transcriptional signatures of ovarian tumours that had high numbers of tumour-infiltrating lymphocytes compared with those that had low numbers<sup>23</sup>. Even within a given lesion, the distribution of immune cell infiltration is not uniform. When areas of cell clustering are present on the leading edge of the lesion this has a different prognostic significance to central areas of clustering, indicating the crucial nature of intratumoural localization<sup>22</sup>.

T-cell activation involves both stimulatory and inhibitory checkpoint signals to finely tune responses to prevent excessive damage and autoimmunity. A direct means of usurping cytotoxic T-cell activation in tumours is through continuous engagement of inhibitory receptors on T-cells, such as cytotoxic T-lymphocyte-associated antigen-4 (CTLA-4) and programmed death 1 (PD1) by upregulation of their ligands<sup>24</sup>. An indirect way of preventing antitumour T-cell responses is by the generation of an immunosuppressive environment. Expansion of a myeloid-derived suppressor cell population (MDSC) — collectively referencing neutrophils, immature dendritic cells, monocytes and early myeloid progenitors — upon tumour implantation implicates tumour-initiated endocrine communication with the immune system early on<sup>25</sup> through tumour and CAF secretion of chemokines (for example, granulocyte-macrophage colony stimulating factor, GM-CSF; or granulocyte colony stimulating factor, G-CSF)<sup>26–28</sup>. Recruitment of immunosuppressive myeloid lineages to the tumour not only suppresses adaptive immunity, but also fosters angiogenesis through the secretion of VEGFA, basic fibroblast growth factor (bFGF) and transforming growth factor  $\beta$  (TGF- $\beta$ )<sup>29</sup>. MDSCs also inhibit natural killer cell function and expand the immunosuppressive regulatory T-cell population. In addition, MDSCs can directly inhibit effector T-cell expansion, activation and migration by crucially altering the environment<sup>30</sup>. B cells have been shown to both suppress and support T-cell function, resulting in differential effects on tumorigenesis<sup>31</sup>. Independent of T-cell function, B cells have also been reported to promote tumour progression by enhancing pro-tumoral inflammation<sup>32,33</sup>. Mast cell recruitment has been implicated in tumorigenesis and angiogenesis<sup>34,35</sup>. Tumour-associated macrophages (TAMs) can also drastically affect tumour progression depending on their polarization<sup>36</sup>.

Immune cell subtypes have demonstrated differential prognostic value depending on the indication. For example, in melanoma, colorectal and breast cancer, a strong correlation between high T-cell occupancy and a good clinical outcome has been reported<sup>22</sup>. Other T-cell subsets, such as regulatory T cells, and the helper T cells T<sub>H</sub>1 and T<sub>H</sub>17 cells are associated with poor or good prognosis depending on the type



**Figure 1 | Origins and influence of tumour heterogeneity.** Tumour formation involves the co-evolution of neoplastic cells together with extracellular matrix and vascular endothelial, stromal and immune cells. The tumour niche is a dynamic physical topography in which structural support, access to growth factors, vascular supply and immune cell interactions can vary drastically even within the same lesion. The immune infiltrate can include multiple cell types, these cell populations can have both pro- and anti-tumour functions and can vary in their activation status and their localization within the tumour. The vascular network can differ in regard to the vessel's tissue of origin, maturity (extent of pericyte coverage), interstitial pressure and functionality. Cancer-associated fibroblasts can have significant plasticity and diverge with regard to activation status, localization within the tissue, stress response and origin.

of cancer<sup>22,37,38</sup>. In breast cancer, assessment of pro- and antitumour immunity revealed that high macrophage content relative to cytotoxic T-cell content was correlated with a worse outcome<sup>39</sup> than those with a low content. Interestingly, a stromal signature derived from microdissected breast tumour biopsies predicted outcome independently of tumour subtype<sup>37</sup>. Analysis of these micro-environmental signatures showed that angiogenesis, hypoxia and macrophage-mediated immune suppression signatures were associated with poor outcome. In addition, B-cell content has been associated with a good prognosis in ovarian<sup>40</sup>, as well as breast<sup>41</sup> cancer. Together, such studies provide a strong argument for the consideration of the immune contexture as a powerful contributor to tumour development.

### Resistance to therapies mediated by tumour stroma

In recent years, medical oncology has focused heavily on personalizing therapeutic approaches with the aim of identifying patient subpopulations that would benefit from specific targeted therapeutics. For example, metastatic melanoma patients who harbour *BRAF*(V600E) mutations and received treatment with the *BRAF*-mutant-specific inhibitor vemurafenib demonstrated marked responses<sup>42</sup>. As effective as these compounds are, however, resistance does ultimately develop<sup>43</sup>.

The exploration of therapeutic resistance has largely focused on the tumour cell. Most documented resistance mechanisms involve secondary pathway mutations or bypass mechanisms within the tumour cells, such as *EGFR*(T790M)<sup>44</sup> mutations or *MET* receptor amplification<sup>45,46</sup> in patients with *EGFR*-mutant NSCLC treated with *EGFR* inhibitors and *NRAS* mutations acquired in patients with *BRAF*-mutant melanoma treated with *BRAF* inhibitors<sup>47</sup>. However, the recent identification of mechanisms of therapeutic resistance that were conferred largely by alterations, not in the tumour cells, but in their environment, indicates the importance of understanding the tumour cell extrinsic compartments<sup>48</sup>.



### Fibroblast-mediated resistance

Early co-culture experiments demonstrated that damaged or irradiated fibroblasts could better support tumour cell growth than non-irradiated fibroblasts, suggesting that within a solid tumour, fibroblasts are not passive elements and could potentially respond and affect therapy<sup>49–51</sup>. In addition to the cell-autonomous mechanisms already described, recent work has demonstrated that ligand-dependent activation of receptor tyrosine kinases can lead to resistance. Stromal derived HGF has been shown to render tumour cells resistant to BRAF inhibition in cell lines that harbour the *BRAF*(V600E) mutation<sup>52,53</sup> through increased phosphorylation of its cognate receptor, MET. Moreover, abundance of HGF expression in patients significantly correlated with reduced responsiveness to drug treatment<sup>52</sup>. It remains to be seen whether combinatorial targeting of HGF or MET, together with an inhibitor of *BRAF*(V600E) will lead to more durable responses in patients with *BRAF*(V600E) mutation. In a similar example, elevation of pro-angiogenic platelet-derived growth factor C (PDGF-C) expression by CAFs could mediate refractoriness to anti-angiogenic therapy *in vivo*<sup>54</sup>.

These examples illustrate the importance of assessing potential stromal mediators of inherent resistance and highlight the challenge of elucidating how therapeutic treatment could elicit resistance through unforeseen stromal changes. In a preclinical model of genotoxic injury, investigators identified secreted factors from normal human fibroblasts and showed that WNT-16b could crucially limit tumour response through paracrine signalling. Increased levels of ligand enhanced tumour cell proliferation and promoted a mesenchymal phenotype in a WNT-dependent manner. Depletion of WNT-16b specifically from fibroblasts enhanced chemotherapy response<sup>55</sup>. Secretion of WNT-16b from stromal fibroblasts depended on an NF- $\kappa$ B mediated pathway known to mediate stress and inflammatory responses. The response of the supporting stroma to treatment may show a more complicated picture in which stress-response programs in these cells may be limiting treatment efficacy by providing a protective environment for tumour cells.

### Vascular-mediated resistance

It is speculated that tumour vasculature serves as a barrier to optimal drug delivery<sup>56</sup>. Tumour vasculature can be constrained by the dense nature of the tumour stroma, leading to compression of tumour vessels, disruption of efficient blood flow and elevation of interstitial pressure, all of which may obstruct movement within and across tumour vessels. Recent work suggests that cytorreduction of the stroma, in this case through enzymatic destruction of hyaluronan, could reduce interstitial pressure and improve vessel patency and flow<sup>57</sup>. The implication is that drug delivery would theoretically be improved, as indicated by the increased efficacy of standard-of-care chemotherapy when combined with hyaluronan depletion in an animal model of pancreatic cancer<sup>57</sup>.

Normalization of leaky vascular beds within tumours through VEGFA pathway inhibition has also been proposed to transiently increase drug delivery in solid tumours<sup>56</sup>. However, the clinical data pertaining to this hypothesis are mixed. For example, when co-administered with anti-VEGF therapy, delivery of radiolabelled chemotherapy was slightly decreased within the tumours of NSCLC patients<sup>58</sup> monitored using positron emission tomography (PET). The biological consequence of these delivery changes is unclear because patients still derive benefit from combination treatments, but changing the delivery may have implications for dosing sequence — this is still under investigation.

Recent work has highlighted a direct role for endothelial cells in tumour response to therapies through secreted factors. Treatment with the chemotherapy drug doxorubicin induced thymic endothelial cell production of the growth factors interleukin 6 (IL-6) and tissue inhibitor of metalloproteinase 1 (TIMP-1), rendering the thymus a protective reservoir for tumour cells during treatment<sup>59</sup>. The concept that a physical niche protects tumour cells during drug treatment has also been suggested for the perivascular space within tumours. The observation of tumour re-initiating cells along tumour vessels<sup>60,61</sup> suggests these locations have a protective role. Paracrine signalling from endothelial cells within this niche has been

shown to increase chemoresistance by inducing a stem-cell-like phenotype in a subset of colorectal tumour cells<sup>62</sup>. Similarly, hypoxic regions of tumours can harbour and support the survival of colon cancer stem cells during chemotherapy<sup>63</sup>. Studies such as these suggest that distinct niches within a tumour could support and instruct tumour regrowth following treatment. These insights highlight the conundrum of the unanticipated effects that therapeutic interventions can have on non-tumour cell components, which can then limit treatment efficacy.

### Immune-mediated resistance

As already mentioned, characterization of immune infiltrate has generated a great deal of interest as data demonstrate that tumour immune contexture may be used to predict patient outcome. This observation suggests that the immune system is an active component of the disease and could affect clinical response and resistance.

In the context of anti-angiogenic therapy, tumours may be rendered refractory to anti-VEGF therapy by a pro-inflammatory micro-environment that includes multiple cell types such as immature myeloid cells and TAMs that secrete factors compensating for VEGF loss to support angiogenesis<sup>64</sup>. Depletion of MDSC expansion and recruitment that is mediated predominantly by secretion of G-CSF in anti-VEGF insensitive experimental models could rescue responsiveness to VEGF depletion, leading to decreased vessel density and tumour growth<sup>26,65</sup>.

Similarly, recent work has suggested that macrophage abundance may impede therapeutic responses. After radiation treatment, levels of factors that mediate macrophage trafficking were significantly increased in the serum of patients with prostate cancer<sup>66</sup>. In patients with node-positive breast cancer, who had undergone intense chemotherapy, those harbouring tumours with a high macrophage, high CD4, but low CD8 T-cell signature had significantly reduced recurrence-free survival<sup>39</sup> compared with patients who had a low macrophage, low CD4 and high CD8 signature. Experimentally, cytotoxic chemotherapy or radiation treatment increased factors that mediate macrophage trafficking and intratumoral macrophages. By preventing treatment-induced macrophage recruitment through blockade of the cognate macrophage receptor, colony stimulating factor receptor-1 responses were improved. Together, these results suggest that clinical responses to chemotherapy or radiotherapy in breast and prostate cancer may be improved by preventing treatment-induced recruitment of suppressive macrophages.

In pre- and post-treatment biopsies from patients with melanoma who received a BRAF inhibitor alone or in combination with an MEK inhibitor, there was an increase in tumour antigen expression that correlated with increased infiltration of cells that carry the CD8 antigen (CD8<sup>+</sup>). This result suggests a positive contribution of cytotoxic T cells to the therapeutic effect. Interestingly, when patients were re-biopsied at time of progression, the CD8<sup>+</sup> infiltrate had decreased concomitant with the emergence of markers of immune cell exhaustion and T-cell inhibitory ligands<sup>67</sup>. This example implicates adaptive immunity in both efficacy and resistance to targeted therapeutics.

Therapeutically harnessing the anti-tumour effects of adaptive immunity in patients has been clearly demonstrated through adoptive cell transfer (ACT). The aim of ACT is to boost a patient's anticancer immunity by transplanting T cells that recognize tumour-specific antigens, leading to recognition and elimination<sup>68</sup>. Although powerful, the responses are not always sustained. Recent work in an experimental model of melanoma suggests that inflammation resulting from the initial tumour response leads to environmental changes that induce loss of the targeted tumour antigens. The presence of tumour necrosis factor- $\alpha$  (TNF- $\alpha$ ) secreted by infiltrating macrophages was identified as the key contributor to the change in target expression by tumour cells<sup>69</sup>.

In summary, the immune system can be implicated in both inherent, as well as acquired resistance to targeted therapies. Even in cases in which the immune cells are actively driving the initial responses to targeted therapies, a wide array of immune suppressive mechanisms may eventually evolve, ultimately leading to tumour progression.

**Table 1 | Examples of therapies that target the tumour stroma listed by compartment**

Molecule	Target	Molecule type	Company	Status (reference)
<b>ECM/fibroblasts</b>				
Marimastat	MMP — broad spectrum	Small molecule	British Biotech	Phase III negative for NSCLC, SCLC and breast cancer (NCT00002911, NCT00003010, NCT00003011)
Prinomastat	MMP 2, 3, 9, 13 and 14	Small molecule	Agouron/Pfizer	Phase III negative for NSCLC and prostate cancer (NCT00004199, NCT00003343)
Tanomastat	MMP 2, 3 and 9	Small molecule	Bayer	Phase III terminated (NCIC-CTG trial OV12)
Neovastat	VEGFR2, MMP 2, 9, 12	Small molecule	AEterna Laboratories	Phase III negative for NSCLC (NCT00005838)
Rebimastat	MMP 1, 2, 8, 9 and 14	Small molecule	Bristol-Myers Squibb	Phase III negative for NSCLC (NCT00006229)
Vismodegib	SMO	Small molecule	Genentech/Roche	Phase II negative for CRC and ovarian cancer and phase II for PDAC (NCT00636610, NCT00739661, NCT01064622)
Saridegib	SMO	Small molecule	Infinity Pharmaceuticals	Phase II terminated for PDAC (NCT01130142, NCT01310816)
Sonidegib	SMO	Small molecule	Novartis	Phase III (NCT01708174)
<b>Vasculature</b>				
Bevacizumab	VEGFA	Antibody	Genentech/Roche	FDA-approved ((BLA) 125085)
Vandetanib	VEGFRs, PDGFRs, EGFR	Small molecule	AstraZeneca	FDA-approved ((NDA) 022405)
Sunitinib	VEGFRs, PDGFRs, FLT3, CSF1R	Small molecule	Pfizer	FDA-approved ((NDA) 021938)
Axitinib	VEGFRs, PDGFRs, KIT	Small molecule	Pfizer	FDA-approved ((NDA) 202324)
Sorafenib	VEGFRs, RAF, PDGFRs, KIT	Small molecule	Bayer	FDA-approved ((NDA) 021923)
Pazopanib	VEGFRs, PDGFRs, KIT	Small molecule	GlaxoSmithKline	FDA-approved ((NDA) 022465)
Cabozantinib	VEGFR2, RET, MET	Small molecule	Exelixis	FDA-approved ((NDA) 203756)
Ziv-aflibercept	VEGFA, VEGFB, PlGF	Receptor–Fc fusion	Regeneron	FDA-approved ((BLA) 125418)
Cilengitide	Integrins $\alpha_3\beta_3$ , $\alpha_5\beta_5$ and $\alpha_5\beta_1$	Small molecule	Merck Serono	Phase III negative for GBM (NCT00689221)
AMG386	ANG2	RP–Fc fusion protein	Amgen	Phase III (NCT01281254)
Parsatuzumab	EGFL-7	Antibody	Genentech/Roche	Phase II (NCT01399684, NCT01366131)
Enoticumab	DLL4	Antibody	Regeneron	Phase I (NCT00871559)
Demcizumab	DLL4	Antibody	OncoMed	Phase I (NCT00744562, NCT01189968, NCT01189942, NCT01189929)
Nesvacumab	ANG2	Antibody	Regeneron	Phase I (NCT01688960, NCT01271972)
<b>Immune</b>				
Ipilimumab	CTLA-4	Antibody	Bristol-Myers Squibb	FDA-approved ((BLA) 125377)
Sipuleucel-T	PAP	DC vaccine	Dendreon	FDA-approved ((BLA) 125197)
Aldesleukin	IL-2	RP	Prometheus	FDA-approved ((BLA) 103293)
IFN $\alpha$ -2b	IFN- $\alpha$ receptor	RP	Merck	FDA-approved ((BLA) 103132)
MK-3475	PD1	Antibody	Merck	Phase III (NCT01866319)
Nivolumab	PD1	Antibody	Bristol-Myers Squibb	Phase III (NCT01668784, NCT01673867, NCT01642004, NCT01721772, NCT01721746, NCT01844505)
Nivolumab	OX40	Antibody	Bristol-Myers Squibb and PPMC	Phase III (NCT01668784, NCT01642004, NCT01673867, NCT01721772, NCT01721746, NCT01844505)
MPDL3280A	PDL1	Antibody	Genentech/Roche	Phase II (NCT01846416)
PLX3397	KIT, CSF1R, FLT3	Small molecule	Plexxikon	Phase II (NCT01349036)
BMS-663513	CD137 (4-1BB)	Antibody	Bristol-Myers Squibb	Phase II (NCT00612664)
Blinatumomab	CD3 and CD19	Bi-specific scFv	Amgen	Phase II (NCT01741792, NCT01466179, NCT01207388, NCT01471782, NCT00560794, NCT01209286)
AMG 820	CSF1R	Antibody	Amgen	Phase I (NCT01444404)
AMP-224	PD1	Antibody	GlaxoSmithKline	Phase I (NCT01352884)
TRX-518	GITR	Antibody	GITR, Inc.	Phase I (NCT01239134)
IMC-CS4	CSF1R	Antibody	ImClone/Eli Lilly	Phase I (NCT01346358)
CP-870,893	CD40	Antibody	Pfizer	Phase I (NCT00711191, NCT01008527, NCT00607048, NCT01456585, NCT01103635)

References listed pertain to the molecule as a stromal-modifying agent, either the FDA application, where approved, or the national clinical trial identification of the oncology trial in the latest phase is listed (note that in some cases the drug may also be tested or approved for an indication for which it acts directly on the tumour cell compartment, which will not be referenced here). ANG2, angiopoietin 2; BLA, biological license application; CRC, colorectal cancer; CSF1R, colony stimulating factor 1 receptor; CTLA-4, cytotoxic T-lymphocyte-associated antigen 4; DC, dendritic cell; DLL4, delta-like 4; ECM, extracellular matrix; EGFL-7, epidermal growth factor like 7; EGFR, epidermal growth factor receptor; FDA, Food and Drug Administration; FLT3, Fms-like tyrosine kinase 3; GBM, glioblastoma multiforme; GITR, glucocorticoid-induced TNFR-related; IFN, interferon; IL-2, interleukin 2; MMP, matrix metalloproteinase; NCT, national clinical trial; NDA, new drug application; NSCLC, non-small cell lung cancer; PAP, prostatic acid phosphatase; PD-1, programmed death-1; PDAC, pancreatic ductal adenocarcinoma; PDGFR, Platelet-derived growth factor receptor; PDL1, programmed death ligand 1; PLGF, placental growth factor; PPMC, Portland Providence Medical Center; RP, recombinant peptide; scFv, single-chain Fv; SCLC, small-cell lung cancer; SMO, smoothened; VEGF, vascular endothelial growth factor; VEGFR, vascular endothelial growth factor receptor.



## Targeting the tumour micro-environment in the clinic

An adaptive, continuous dialogue exists between tumour cells and their surroundings, mediated through direct cell contact with stromal components or through secreted signalling factors (cytokines, chemokines and growth factors). As already discussed, the tumour niche shapes responses to the selective pressure of drug treatment and may affect the emergence of resistance. Can durable responses result from multi-pronged approaches that target both the tumour cells and their cell-extrinsic support? To envisage a scenario in which targeted therapeutics for multiple compartments are successfully implemented, it is important to review what has been developed and clinically tested for targeting the tumour stroma (Table 1).

### Targeting stromal fibroblasts

The impressive preclinical data and the potential broad application of matrix-metalloproteinase inhibitors for oncology has fuelled clinical development and testing of several pan-protease inhibitors, including tanomastat, marimastat and prinomastat. These agents failed to show significant benefit over the standard-of-care treatment across multiple forms of cancer<sup>70</sup>. Retrospectively, several explanations have been posited to explain these outcomes. Although matrix metalloproteinases are almost ubiquitously overexpressed in human tumours, it is likely that early-stage tumours are more dependent on their activity than late-stage, established tumours. Moreover, our incomplete understanding of matrix metalloproteinase biology and expression in advanced carcinomas at that time could explain some of the failures. For example, tanomastat does not target the crucial matrix metalloproteinases expressed that correlate with poor prognosis in small-cell lung cancer<sup>70,71</sup>.

Targeting signalling pathways that affect tumour mesenchyme have also yielded disappointing results thus far. As is the case for the Hedgehog signalling pathway in which extensive preclinical data showed increased expression of Hedgehog ligands in many tumour types. Hedgehog ligands released by tumour cells lead to activation of the pathway in stromal fibroblasts<sup>72</sup>, inducing the production of secreted factors that contribute to tumorigenesis by acting on tumour cells or other stromal compartments such as the vasculature. However, in contrast to tumours in which the Hedgehog pathway is mutated, the inhibitors of the protein smoothened, vismodegib and saridegib, have so far failed to show a benefit in colorectal<sup>73</sup>, ovarian<sup>74</sup> or pancreatic<sup>75,76</sup> cancer when combined with standard-of-care therapies. It has been postulated that Hedgehog pathway inhibition would reduce the desmoplastic stroma within pancreatic tumours and lead to increased vessel formation, thereby facilitating drug delivery<sup>77</sup>. However, in preclinical models, this effect led only to very small and transient responses that did not translate in clinical trials — in which combined treatments of Hedgehog pathway inhibitors with chemotherapy have not shown clinical benefit compared with chemotherapy alone<sup>75,76</sup>.

### Targeting vasculature

The development of molecules that target VEGFA pathways represents the most successful approach so far in targeting the tumour environment. Bevacizumab, a humanized, monoclonal anti-VEGFA antibody was the first molecule targeting this pathway to be approved by the US Food and Drug Administration (FDA)<sup>78</sup>. In combination with chemotherapy, bevacizumab provides a benefit to patients with advanced NSCLC<sup>79</sup> and metastatic colorectal cancer<sup>80</sup>. A benefit was also demonstrated for metastatic renal cancer<sup>81</sup> when the antibody was combined with interferon- $\alpha$  and recurrent glioblastoma multiforme (GBM)<sup>82,83</sup> as monotherapy. More recently, improvements in progression-free survival were reported in patients with high-risk ovarian cancer<sup>84</sup> and in overall survival for patients with advanced metastatic cervical cancer<sup>85</sup>. The FDA has since approved several small molecule inhibitors, including sunitinib, axitinib, pazopanib, vandetanib, cabozantinib and sorafenib, that target the VEGF receptor VEGFR2 and other receptor tyrosine kinases for multiple cancers.

Although all clinical trials have demonstrated some degree of activity for bevacizumab in combination with chemotherapy, the extent varies depending on the cancer type, with only marginal activity seen in pancreatic cancer<sup>86</sup> for example. Although it is unclear what drives heterogeneity

in sensitivity across cancer types, it is plausible that it could be due to inherent differences in vascular maturity, expression of compensating pro-angiogenic factors and differential dependency of tumour cells on oxygen and nutrients. Pancreatic tumours are noted for their hypovascularity<sup>77</sup>, implying a lesser dependency on vascular supply than other tumour types.

Recent preclinical data have suggested that small molecule anti-angiogenic therapy could increase invasiveness and metastasis<sup>87,88</sup>. Increased metastasis was confirmed with a VEGFR tyrosine kinase inhibitor (TKI) at supraclinical dose levels, but not with anti-VEGF monoclonal antibodies in multiple autochthonous animal models<sup>89</sup>, consistent with clinical data in many epithelial cancers, including metastatic breast, kidney, colorectal and pancreatic cancers<sup>90,87</sup>. Newly evaluated clinical data from metastatic renal cell carcinoma failed to show evidence of accelerated disease or a worse outcome with the VEGFR TKI sunitinib<sup>91</sup>. These reports highlight the need to distinguish between different classes of therapeutic that target the same pathway and emphasize the importance of accurate preclinical modelling of clinical dosing regimens<sup>92</sup>. GBM may be an exception for which clinical data support the possibility that inhibiting VEGF in a subset of patients may exacerbate the invasive phenotype of their tumours<sup>93</sup>. Recent preclinical data have provided a context-specific mechanism in GBM, whereby VEGF functions as a negative regulator of invasive signalling pathways driven by the MET receptor that is expressed in tumour cells<sup>94</sup>, suggesting that combination therapy with MET inhibitors may be worth exploring.

### Targeting immune cells

The recent approval of several immune modulatory therapies for the treatment of cancer has reinvigorated the field of tumour immunotherapy<sup>95</sup>. In 2010, the FDA approved a cell-based therapy called Provenge (sipuleucel-T) for the treatment of castration-resistant prostate cancer. Although the clinical results showed no tumour regression, overall survival was increased by more than 4 months<sup>96</sup>. This was followed shortly thereafter by the approval of ipilimumab, a monoclonal antibody targeting the negative immune checkpoint protein CTLA-4 (ref. 97). Patients with resistant, refractory melanoma demonstrated a two-fold increase in survival when treated with ipilimumab. In a follow-up trial, ipilimumab also showed significant efficacy in naive melanoma patients together with chemotherapy, when compared with chemotherapy alone<sup>98</sup>. By blocking CTLA-4, ipilimumab allows for enhanced antitumour effector T-cell function and is also thought to inhibit immunosuppressive regulatory T-cell function. Based on these clinical successes, there is renewed interest in the clinical development of therapies that either block immunosuppressive mechanisms, such as PD1 and its ligand PDL1, to restore T-cell function or enhance immune function by engaging co-stimulatory receptors such as OX40 with agonist antibodies.

Other technologies that are currently in clinical development attempt to directly engage T-cell-mediated killing. Early clinical experience of blinatumomab, a bispecific antibody fragment with dual affinity for CD3 on T cells and CD19 on B cells, in patients with non-Hodgkin's lymphoma is promising because exceptionally low doses of this agent led to tumour regression<sup>99,100</sup>. Such direct tumour targeting of T cells is attractive in that it provides specificity to T-cell-mediated lysis of tumour cells and may avoid most immune escape mechanisms.

Targeting components of the innate immune system is also under clinical evaluation. For example, an anti-CD40 agonist antibody has shown early clinical promise in combination with gemcitabine in patients with pancreatic cancer<sup>101</sup>. The mechanism underlying these responses has been attributed to macrophage infiltration, rather than enhanced antigen-presenting cell function and T-cell activity.

### Therapeutic combinations

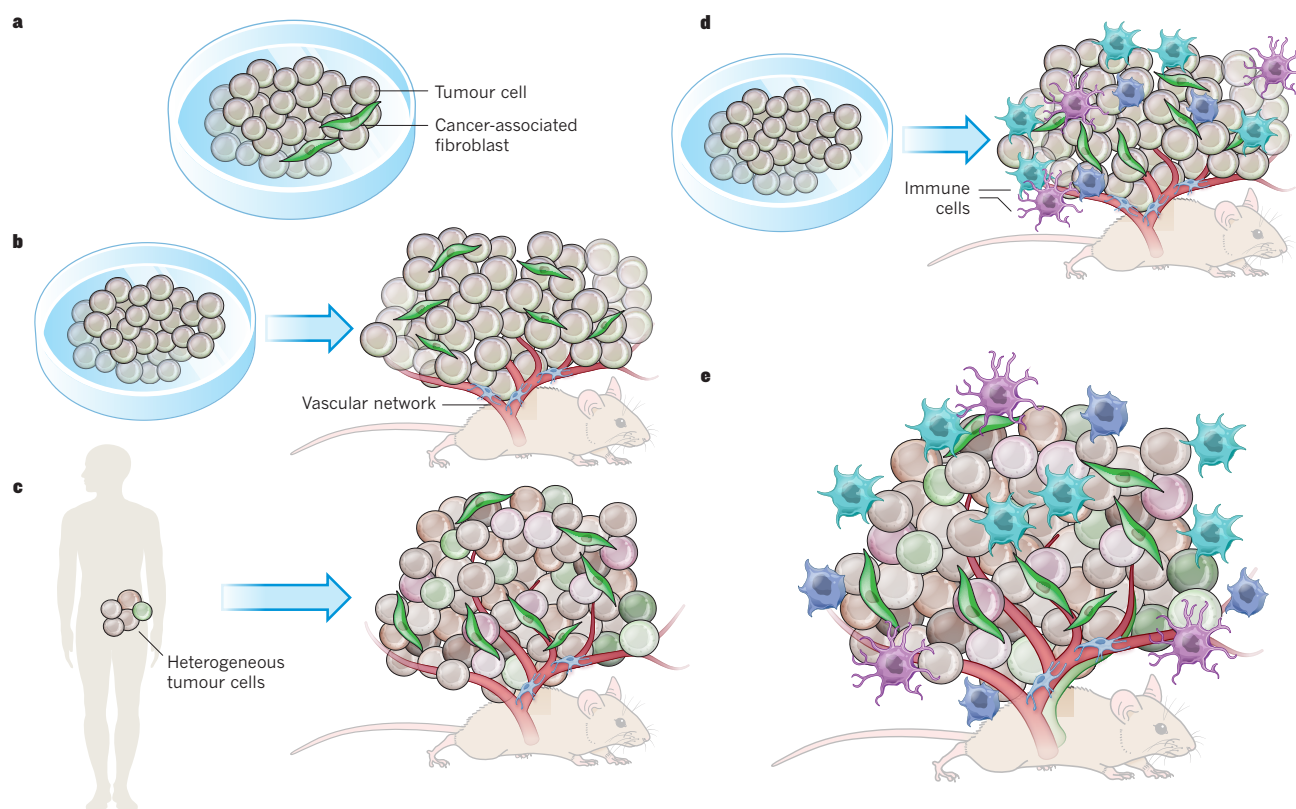
Given the growing availability of agents targeting different tumour compartments, it seems promising that combination therapy may lead to more robust or durable responses. The fundamental question is how to combine

## BOX 1

## Preclinical model systems

Accurately modelling tumour complexity and heterogeneity poses a significant challenge. Moreover, assessing efficacy preclinically when targeting multiple compartments of the tumour is further complicated by the absence of distinct tumour compartments in some preclinical model systems. Such limitations mean that no single model will be sufficient for comprehensive preclinical evaluation. The throughput of *in vitro* tumour-cell culture (Fig. **a**) is robust, and cytokines, chemokines and growth factors can all be added to assess impact on tumour cell responsiveness, or they can be co-cultured with stromal cells. However the results can be difficult to interpret and components of the tumour micro-environment, such as vascular and immune cells are missing. Xenograft studies with human cell lines (Fig. **b**) can incorporate features of the tumour micro-environment such as vascular and fibroblasts but these cells are of mouse origin (in contrast to the tumour cells, which are of human origin) and immune cells are absent. In addition, these studies do not account for tumour heterogeneity.

Patient-derived xenografts studies with human tumour explants (Fig. **c**) can incorporate features of the tumour micro-environment and are representative of the original tumour cell heterogeneity; however, the host is immunodeficient and, in the same way as studies with cell lines, the components are of mouse origin whereas the tumour cells are of human origin. In addition to incorporating features of the micro-environment such as vascular and stromal cells, syngeneic studies (Fig. **d**) offer the possibility of an immune component; however, limited murine cell lines are available, questions exist regarding the immunogenicity of these lines and injection may induce inflammation. Autochthonous murine tumour models (Fig. **e**) incorporate many features of the tumour and the tumour micro-environment. The tumour initiated within the native tissue context and evolves with all the heterogeneous stromal cell types but has a reduced mutational spectrum. However, the throughput is low and the experimentation can be expensive.



these agents clinically. Clearly, a deeper understanding of their mechanisms of action will be required to guide rational drug combinations.

Chemotherapy may pose a challenge for combination treatment with stromal targeting strategies. Although bevacizumab is tolerated and effective when combined with standard-of-care chemotherapy regimens across multiple indications, it is uncertain whether chemotherapy and immune modulatory therapies will be combined successfully. Chemotherapy could potentially enhance efficacy of immunotherapies through multiple mechanisms, such as increased antigen production, improved antigen presentation, augmented T-cell response and trafficking<sup>102</sup>. However, chemotherapy can antagonize immune modulatory effects<sup>103</sup>. In addition to certain chemotherapies leading to neutropaenia,

the type and context of cell death induced may be crucial for the success of combinations. The requirement for adaptive immunity in mediating chemotherapy-induced responses is still unclear, as divergent results have been reported using independent model systems<sup>104,105</sup>. This highlights the impact of preclinical model choice on experimental outcome. Although it is unclear how researchers would predict effective combinations of therapy for certain tumours, recent clinical success of immunotherapy-chemotherapy combinations, for example ipilimumab with dacarbazine<sup>98</sup> for the treatment of metastatic melanoma and vaccination with docetaxel for metastatic androgen-resistant prostate cancer<sup>106</sup>, should force the further consideration of chemotherapy and immunotherapy combinations.

The role of the endothelium in regulating immune cell trafficking



suggests there is an intimacy between endothelial cell modulation and immunotherapy that is worth exploring. Strategies that increase vessel patency have been linked to increased immune cell infiltration into tumours<sup>107,108</sup>. In experiments, otherwise inert tumour vaccination was found to be rendered efficacious, possibly through modulation of the endothelium allowing enhanced immune cell passage and tumour access<sup>23,107</sup>. Furthermore, by using adaptive transfer techniques VEGF blockade could increase intratumoral infiltration of immune cells<sup>109</sup>. It has also been postulated that anti-angiogenic treatment could enhance dendritic cell function, thereby enhancing immunotherapy responses<sup>110</sup>. The rationale for anti-angiogenic and immunotherapeutic combinations is compelling and clinical interrogation of such combinations is ongoing.

Combinations of targeted therapies that may synergize are currently being explored. Clinical interrogation of the ipilimumab and the *BRAF*-mutant kinase inhibitor vemurafenib combination is currently being tested in metastatic melanoma. MEK pathway inhibition has been demonstrated to hinder T-cell function, despite the ability of inhibition to increase melanocyte differentiation antigen expression<sup>67,111</sup>. However, this is not the case when the pathway is inhibited as a result of vemurafenib treatment, which does not affect T-cell function<sup>112</sup>. Interestingly, in pancreatic cancer, MEK inhibition has been linked to the reduction of MDSC populations that permits T-cell-mediated tumour targeting<sup>27,28</sup>.

### Future challenges

Despite the success in targeting non-tumour cell compartments, significant challenges still lie ahead for implementing stromal targeting strategies in clinical practice. These range from difficulties in assessing the composition of the stroma in human tumours to correctly modelling preclinically the vast heterogeneity observed in human tumours and correlating this heterogeneity with outcome, drug response and drug resistance.

Accurately modelling tumour stroma complexity and heterogeneity preclinically is a challenge<sup>113</sup> (Box 1). Not least because the complexity is poorly defined. The complexity of the tumour stroma at various stages of development and after drug treatment remains to be characterized in great detail, and this presents its own challenges. Modelling the multiple non-tumour cell types may be particularly complicated, as most preclinical techniques rely on implantation of tumour cells in foreign sites (most often under the skin, where the stromal representation may be different) or in immune compromised hosts that lack crucial immune effector cells. Looking ahead, the preclinical evaluation of combination therapies that target multiple tumour components will require the incorporation of several types of preclinical system and most likely the development of complex genetically engineered models.

Another obstacle in targeting the tumour stroma is developing reliable diagnostic markers that are based on a clear understanding of the determinants of responsiveness. For example, although numerous attempts have been made to identify predictive biomarkers for the effectiveness of bevacizumab, so far, no reliable and reproducible predictive biomarkers have been identified<sup>21,114,115</sup>. Intra- and intertumour heterogeneity may explain why a biomarker has remained elusive for therapies targeting the tumour stroma in general, and for anti-angiogenic therapies in particular. For instance, in metastatic NSCLC, primary lung tumours and their matched brain metastases lack correlation with respect to micro-vessel density, vascular maturity or VEGF expression<sup>116</sup>, illustrating how tumour vasculatures vary drastically even within the same patient. These vascular differences may be reflective of the unique anatomical locations within which the lesions emerged, which is suggestive of micro-environmental influence on tumour growth, but nonetheless confounds the predictability of therapeutic response. Similarly, intra- and intertumour heterogeneity is observed with immune infiltrates, significantly complicating the evaluation of predictive biomarker identification. Additional challenges for immunotherapy lie in determining diagnostic and biomarker criteria for therapies whose targets are particularly heterogeneous or of low abundance in the tumour, as is the case for anti-CTLA-4, in which only a subset of non-tumour cells express the CTLA-4 ligand. In this case as well, despite concerted efforts, no reliable biomarkers have been identified<sup>24</sup>.

These observations highlight the inherent difficulty in the development of predictive biomarkers for therapies that target highly heterogeneous and somewhat discrete compartments in tumours.

In light of immense tumour heterogeneity, it is unclear how and when to clinically evaluate tumour heterogeneity. As discussed, the complexity of the tumour is a result of continuous crosstalk between the tumour cells and the environment over time, adding the complication of temporal heterogeneity on top of spatial heterogeneity. Given that patients with cancer are often treated through multiple lines of therapy, serial biopsies from different sites may be required to predict response and implement therapy. This is particularly relevant for treatment-induced stromal alterations that are not observable from pretreatment biopsies. Although serial biopsies could be incorporated into routine practice as a means to inform and guide each line of treatment, as has recently been done<sup>117</sup>, this poses a significant burden on patients.

Another challenge is how to measure clinical benefit for agents that target the tumour stroma. Patients can derive benefit from bevacizumab after disease progression<sup>118</sup>, suggesting that the traditional definition of disease progression based on Response Evaluation Criteria in Solid Tumors (RECIST) may not be the best method to measure drug benefit when therapeutically targeting the tumour stroma. Immunotherapy strategies have unique challenges in this realm. For example, regulations of response criteria needed to be altered for this class of oncology medications<sup>119</sup> because patients demonstrate delayed regression or even pseudo-progression after treatment.

Finally, understanding and managing stroma-mediated resistance would have a profound affect on therapeutic strategies that target this compartment. The emergence of therapeutic resistance should theoretically be reduced if the target population is genetically stable; targeting the tumour stroma, therefore, could be advantageous because mutations have rarely been identified in the stroma<sup>120–122</sup>. Support for this hypothesis may be gleaned from anti-VEGFA clinical data. Following disease progression in patients on chemotherapy, bevacizumab treatment provided a benefit when given with a new chemotherapy regimen<sup>123</sup>. Moreover, clinical bevacizumab efficacy was evaluated when given during first-line treatment with chemotherapy and then continued with an alternative regimen at progression. The addition of bevacizumab to second-line therapy significantly improved overall and progression-free survival, demonstrating that progression in first-line therapy was not due to anti-VEGF resistance<sup>118</sup>. Despite the proposed advantage provided by genetic stability of the stroma, this tissue will continue to evolve during treatment — and adaptation by non-genetic means may complicate the scenario, especially given that under treatment pressure epigenetic changes within the stroma have been reported<sup>124,125</sup>.

Therapeutic resistance reflects active tumour evolution, and environmental mediated resistance is illustrative of dynamic interplay between tumour cells and their surroundings when the selective pressure of drug therapy is applied. A more comprehensive understanding of heterogeneity within the tumour cell extrinsic compartments and how this influences resistance development will expand our understanding of treatment responses. Furthermore, the elucidation of stromal-mediated mechanisms of resistance will probably lead to the discovery of new therapeutic options. The targeting of multiple tumour compartments may represent a solution to avoid resistance and achieve durable patient responses; however, elucidating the relationship between tumour heterogeneity and therapeutic response is still a significant challenge. ■

Received 17 May; accepted 7 August 2013.

1. Hanahan, D. & Coussens, L. M. Accessories to the crime: functions of cells recruited to the tumor microenvironment. *Cancer Cell* **21**, 309–322 (2012).
2. Egeblad, M., Nakasone, E. S. & Werb, Z. Tumors as organs: complex tissues that interface with the entire organism. *Dev. Cell* **18**, 884–901 (2010).
3. Dotto, G. P., Weinberg, R. A. & Ariza, A. Malignant transformation of mouse primary keratinocytes by Harvey sarcoma virus and its modulation by surrounding normal cells. *Proc. Natl Acad. Sci. USA* **85**, 6389–6393 (1988).
4. Orimo, A. *et al.* Stromal fibroblasts present in invasive human breast carcinomas promote tumor growth and angiogenesis through elevated SDF-1/CXCL12 secretion. *Cell* **121**, 335–348 (2005).

5. Polanska, U. M. & Orimo, A. Carcinoma-associated fibroblasts: non-neoplastic tumour-promoting mesenchymal cells. *J. Cell. Physiol.* **8**, 1651–1657 (2013).
6. Rinn, J. L., Bondre, C., Gladstone, H. B., Brown, P. O. & Chang, H. Y. Anatomic demarcation by positional variation in fibroblast gene expression programs. *PLoS Genet.* **2**, e119 (2006).
7. Rudnick, J. A. *et al.* Functional heterogeneity of breast fibroblasts is defined by a prostaglandin secretory phenotype that promotes expansion of cancer-stem like cells. *PLoS ONE* **6**, e24605 (2011).
8. Bergers, G. *et al.* Matrix metalloproteinase-9 triggers the angiogenic switch during carcinogenesis. *Nature Cell Biol.* **2**, 737–744 (2000).
9. Quante, M. *et al.* Bone marrow-derived myofibroblasts contribute to the mesenchymal stem cell niche and promote tumor growth. *Cancer Cell* **19**, 257–272 (2011).
10. Yamashita, M. *et al.* Role of stromal myofibroblasts in invasive breast cancer: stromal expression of alpha-smooth muscle actin correlates with worse clinical outcome. *Breast Cancer* **19**, 170–176 (2012).
11. Fujita, H. *et al.*  $\alpha$ -Smooth muscle actin expressing stroma promotes an aggressive tumor biology in pancreatic ductal adenocarcinoma. *Pancreas* **39**, 1254–1262 (2010).
12. Vihinen, P. & Kähäri, V.-M. Matrix metalloproteinases in cancer: prognostic markers and therapeutic targets. *Int. J. Cancer* **99**, 157–166 (2002).
13. Calle, E. E. & Kaaks, R. R. Overweight, obesity and cancer: epidemiological evidence and proposed mechanisms. *Nature Rev. Cancer* **4**, 579–591 (2004).
14. Ferrara, N. & Kerbel, R. S. Angiogenesis as a therapeutic target. *Nature* **438**, 967–974 (2005).
15. Kalluri, R. Basement membranes: structure, assembly and role in tumour angiogenesis. *Nature Rev. Cancer* **3**, 422–433 (2003).
16. Rosenberg, R. D. & Aird, W. C. Vascular-bed-specific hemostasis and hypercoagulable states. *N. Engl. J. Med.* **340**, 1555–1564 (1999).
17. Trédan, O., Galmarini, C. M., Patel, K. & Tannock, I. F. Drug resistance and the solid tumor microenvironment. *J. Natl Cancer Inst.* **99**, 1441–1454 (2007).
18. Meert, A.-P. *et al.* The role of microvessel density on the survival of patients with lung cancer: a systematic review of the literature with meta-analysis. *Br. J. Cancer* **87**, 694–701 (2002).
19. Des Guetz, G. *et al.* Microvessel density and VEGF expression are prognostic factors in colorectal cancer. Meta-analysis of the literature. *Br. J. Cancer* **94**, 1823–1832 (2006).
20. Uzzan, B., Nicolas, P., Cucherat, M. & Perret, G.-Y. Microvessel density as a prognostic factor in women with breast cancer: a systematic review of the literature and meta-analysis. *Cancer Res.* **64**, 2941–2955 (2004).
21. Hegde, P. S. *et al.* Predictive impact of circulating vascular endothelial growth factor in 4 phase III trials evaluating bevacizumab. *Clin. Cancer Res.* **19**, 929–937 (2013).
22. Fridman, W.-H., Pagès, F., Sautès-Fridman, C. & Galon, J. The immune contexture in human tumours: impact on clinical outcome. *Nature Rev. Cancer* **12**, 298–306 (2012).
23. Buckanovich, R. J. *et al.* Endothelin B receptor mediates the endothelial barrier to T cell homing to tumors and disables immune therapy. *Nature Med.* **14**, 28–36 (2008).
24. Pardoll, D. M. The blockade of immune checkpoints in cancer immunotherapy. *Nature Rev. Cancer* **12**, 252–264 (2012).
25. Yang, L. *et al.* Expansion of myeloid immune suppressor Gr<sup>+</sup>CD11b<sup>+</sup> cells in tumor-bearing host directly promotes tumor angiogenesis. *Cancer Cell* **6**, 409–421 (2004).
26. Shojaei, F. *et al.* G-CSF-initiated myeloid cell mobilization and angiogenesis mediate tumor refractoriness to anti-VEGF therapy in mouse models. *Proc. Natl Acad. Sci. USA* **106**, 6742–6747 (2009).
27. Pylayeva-Gupta, Y., Lee, K. E., Hajdu, C. H., Miller, G. & Bar-Sagi, D. Oncogenic Kras-induced GM-CSF production promotes the development of pancreatic neoplasia. *Cancer Cell* **21**, 836–847 (2012).
28. Bayne, L. J. *et al.* Tumor-derived granulocyte-macrophage colony-stimulating factor regulates myeloid inflammation and T cell immunity in pancreatic cancer. *Cancer Cell* **21**, 822–835 (2012).
29. Motz, G. T. & Coukos, G. The parallel lives of angiogenesis and immunosuppression: cancer and other tales. *Nature Rev. Immunol.* **11**, 702–711 (2011).
30. Gabrilovich, D. I., Ostrand-Rosenberg, S. S. & Bronte, V. V. Coordinated regulation of myeloid cells by tumours. *Nature Rev. Immunol.* **12**, 253–268 (2012).
31. Nelson, B. H. CD20<sup>+</sup> B cells: the other tumor-infiltrating lymphocytes. *J. Immunol.* **185**, 4977–4982 (2010).
32. de Visser, K. E., Korets, L. V. & Coussens, L. M. *De novo* carcinogenesis promoted by chronic inflammation is B lymphocyte dependent. *Cancer Cell* **7**, 411–423 (2005).
33. Ammirante, M., Luo, J.-L., Grivennikov, S., Nedospasov, S. & Karin, M. B-cell-derived lymphotoxin promotes claudin-resistant prostate cancer. *Nature* **464**, 302–305 (2010).
34. Coussens, L. M. *et al.* Inflammatory mast cells up-regulate angiogenesis during squamous epithelial carcinogenesis. *Genes Dev.* **13**, 1382–1397 (1999).
35. Yang, F.-C. *et al.* *Nf1*-dependent tumors require a microenvironment containing *Nf1*<sup>+/+</sup> and c-kit-dependent bone marrow. *Cell* **135**, 437–448 (2008).
36. Mantovani, A. & Sica, A. A. Macrophages, innate immunity and cancer: balance, tolerance, and diversity. *Curr. Opin. Immunol.* **22**, 231–237 (2010).
37. Finak, G. *et al.* Stromal gene expression predicts clinical outcome in breast cancer. *Nature Med.* **14**, 518–527 (2008).
38. Tosolini, M. *et al.* Clinical impact of different classes of infiltrating T cytotoxic and helper cells (Th1, Th2, Th17) in patients with colorectal cancer. *Cancer Res.* **71**, 1263–1271 (2011).
39. DeNardo, D. G. *et al.* Leukocyte complexity predicts breast cancer survival and functionally regulates response to chemotherapy. *Cancer Discov.* **1**, 54–67 (2011).
40. Nielsen, J. S. *et al.* CD20<sup>+</sup> tumor-infiltrating lymphocytes have an atypical CD27<sup>+</sup> memory phenotype and together with CD8<sup>+</sup> T cells promote favorable prognosis in ovarian cancer. *Clin. Cancer Res.* **18**, 3281–3292 (2012).
41. Schmidt, M. *et al.* The humoral immune system has a key prognostic impact in node-negative breast cancer. *Cancer Res.* **68**, 5405–5413 (2008).
42. Chapman, P. B. *et al.* Improved survival with vemurafenib in melanoma with BRAF V600E mutation. *N. Engl. J. Med.* **364**, 2507–2516 (2011).
43. Wagle, N. *et al.* Dissecting therapeutic resistance to RAF inhibition in melanoma by tumor genomic profiling. *J. Clin. Oncol.* **29**, 3085–3096 (2011).
44. Kobayashi, S. *et al.* *EGFR* mutation and resistance of non-small-cell lung cancer to gefitinib. *N. Engl. J. Med.* **352**, 786–792 (2005).
45. Engelman, J. A. *et al.* *MET* amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. *Science* **316**, 1039–1043 (2007).
46. Bean, J. *et al.* *MET* amplification occurs with or without *T790M* mutations in *EGFR* mutant lung tumors with acquired resistance to gefitinib or erlotinib. *Proc. Natl Acad. Sci. USA* **104**, 20932–20937 (2007).
47. Nazarian, R. *et al.* Melanomas acquire resistance to *B-RAF*(V600E) inhibition by RTK or *N-RAS* upregulation. *Nature* **468**, 973–977 (2010).
48. Meads, M. B., Gatenby, R. A. & Dalton, W. S. Environment-mediated drug resistance: a major contributor to minimal residual disease. *Nature Rev. Cancer* **9**, 665–674 (2009).
49. Barcellos-Hoff, M. H. & Ravan, S. A. Irradiated mammary gland stroma promotes the expression of tumorigenic potential by unirradiated epithelial cells. *Cancer Res.* **60**, 1254–1260 (2000).
50. Krtoch, A., Parrinello, S., Lockett, S., Desprez, P. Y. & Campisi, J. Senescent fibroblasts promote epithelial cell growth and tumorigenesis: a link between cancer and aging. *Proc. Natl Acad. Sci. USA* **98**, 12072–12077 (2001).
51. Ohuchida, K. *et al.* Radiation to stromal fibroblasts increases invasiveness of pancreatic cancer cells through tumor-stromal interactions. *Cancer Res.* **64**, 3215–3222 (2004).
52. Straussman, R. *et al.* Tumour micro-environment elicits innate resistance to RAF inhibitors through HGF secretion. *Nature* **487**, 500–504 (2012).
53. Wilson, T. R. *et al.* Widespread potential for growth-factor-driven resistance to anticancer kinase inhibitors. *Nature* **487**, 505–509 (2012).
54. Crawford, Y. *et al.* PDGF-C mediates the angiogenic and tumorigenic properties of fibroblasts associated with tumors refractory to anti-VEGF treatment. *Cancer Cell* **15**, 21–34 (2009).
55. Sun, Y. *et al.* Treatment-induced damage to the tumor microenvironment promotes prostate cancer therapy resistance through WNT16B. *Nature Med.* **18**, 1359–1368 (2012).
56. Jain, R. K. Normalization of tumor vasculature: an emerging concept in antiangiogenic therapy. *Science* **307**, 58–62 (2005).
57. Provenzano, P. P. *et al.* Enzymatic targeting of the stroma ablates physical barriers to treatment of pancreatic ductal adenocarcinoma. *Cancer Cell* **21**, 418–429 (2012).
58. Van der Veldt, A. A. M. *et al.* Rapid decrease in delivery of chemotherapy to tumors after anti-VEGF therapy: implications for scheduling of anti-angiogenic drugs. *Cancer Cell* **21**, 82–91 (2012).
59. Gilbert, L. A. & Hemann, M. T. DNA damage-mediated induction of a chemoresistant niche. *Cell* **143**, 355–366 (2010).
60. Calabrese, C. *et al.* A perivascular niche for brain tumor stem cells. *Cancer Cell* **11**, 69–82 (2007).
61. Krishnamurthy, S. *et al.* Endothelial cell-initiated signaling promotes the survival and self-renewal of cancer stem cells. *Cancer Res.* **70**, 9969–9978 (2010).
62. Lu, J. *et al.* Endothelial cells promote the colorectal cancer stem cell phenotype through a soluble form of Jagged-1. *Cancer Cell* **23**, 171–185 (2013).
63. Mao, Q. *et al.* A tumor hypoxic niche protects human colon cancer stem cells from chemotherapy. *J. Cancer Res. Clin. Oncol.* **139**, 211–222 (2013).
64. Shojaei, F. *et al.* Tumor refractoriness to anti-VEGF treatment is mediated by CD11b<sup>+</sup>Gr1<sup>+</sup> myeloid cells. *Nature Biotechnol.* **25**, 911–920 (2007).
65. Phan, V. T. *et al.* Oncogenic RAS pathway activation promotes resistance to anti-VEGF therapy through G-CSF-induced neutrophil recruitment. *Proc. Natl Acad. Sci. USA* **110**, 6079–6084 (2013).
66. Xu, J. *et al.* CSF1R signaling blockade stanches tumor-infiltrating myeloid cells and improves the efficacy of radiotherapy in prostate cancer. *Cancer Res.* **73**, 2782–2794 (2013).
67. Frederick, D. T. *et al.* BRAF inhibition is associated with enhanced melanoma antigen expression and a more favorable tumor microenvironment in patients with metastatic melanoma. *Clin. Cancer Res.* **19**, 1225–1231 (2013).
68. Rosenberg, S. A., Restifo, N. P., Yang, J. C., Morgan, R. A. & Dudley, M. E. Adoptive cell transfer: a clinical path to effective cancer immunotherapy. *Nature Rev. Cancer* **8**, 299–308 (2008).



69. Landsberg, J. *et al.* Melanomas resist T-cell therapy through inflammation-induced reversible dedifferentiation. *Nature* **490**, 412–416 (2012).  
**This report implicates treatment-induced inflammation as the cause of therapeutic resistance against an immunotherapy.**
70. Coussens, L. M., Fingleton, B. & Matrisian, L. M. Matrix metalloproteinase inhibitors and cancer: trials and tribulations. *Science* **295**, 2387–2392 (2002).
71. Michael, M. *et al.* Expression and prognostic significance of metalloproteinases and their tissue inhibitors in patients with small-cell lung cancer. *J. Clin. Oncol.* **17**, 1802–1808 (1999).
72. Theunissen, J.-W. & de Sauvage, F. J. Paracrine Hedgehog signaling in cancer. *Cancer Res.* **69**, 6007–6010 (2009).
73. Berlin, J. *et al.* A randomized phase II trial of vismodegib versus placebo with FOLFOX or FOLFIRI and bevacizumab in patients with previously untreated metastatic colorectal cancer. *Clin. Cancer Res.* **19**, 258–267 (2013).
74. Kaye, S. B. *et al.* A phase II, randomized, placebo-controlled study of vismodegib as maintenance therapy in patients with ovarian cancer in second or third complete remission. *Clin. Cancer Res.* **18**, 6509–6518 (2012).
75. Madden, J. I. *Infinity Reports Update From Phase 2 Study of Saridegib Plus Gemcitabine in Patients with Metastatic Pancreatic Cancer* <http://phx.corporate-ir.net/phoenix.zhtml?c=121941&p=irol-newsArticle&ID=1653550&highlight=> (Infinity Pharmaceuticals, 2012).
76. Catenacci, D. *et al.* A phase IB/randomized phase II study of gemcitabine (G) plus placebo (P) or vismodegib (V), a Hedgehog (Hh) pathway inhibitor, in patients (pts) with metastatic pancreatic cancer: Interim analysis of a University of Chicago phase II consortium study. *J. Clin. Oncol.* **30**, (suppl.), abstr. 4022 (2012).
77. Olive, K. P. *et al.* Inhibition of Hedgehog signaling enhances delivery of chemotherapy in a mouse model of pancreatic cancer. *Science* **324**, 1457–1461 (2009).
78. Ferrara, N., Hillan, K. J., Gerber, H.-P. & Novotny, W. Discovery and development of bevacizumab, an anti-VEGF antibody for treating cancer. *Nature Rev. Drug Discov.* **3**, 391–400 (2004).
79. Sandler, A. *et al.* Paclitaxel-carboplatin alone or with bevacizumab for non-small-cell lung cancer. *N. Engl. J. Med.* **355**, 2542–2550 (2006).
80. Hurwitz, H. I. *et al.* Bevacizumab in combination with fluorouracil and leucovorin: an active regimen for first-line metastatic colorectal cancer. *J. Clin. Oncol.* **23**, 3502–3508 (2005).
81. Yang, J. C. *et al.* A randomized trial of bevacizumab, an anti-vascular endothelial growth factor antibody, for metastatic renal cancer. *N. Engl. J. Med.* **349**, 427–434 (2003).
82. Friedman, H. S. *et al.* Bevacizumab alone and in combination with irinotecan in recurrent glioblastoma. *J. Clin. Oncol.* **27**, 4733–4740 (2009).
83. Vredenburgh, J. J. *et al.* Bevacizumab plus irinotecan in recurrent glioblastoma multiforme. *J. Clin. Oncol.* **25**, 4722–4729 (2007).
84. Perren, T. J. *et al.* A phase 3 trial of bevacizumab in ovarian cancer. *N. Engl. J. Med.* **365**, 2484–2496 (2011).
85. Tewari, K. S. *et al.* Incorporation of bevacizumab in the treatment of recurrent and metastatic cervical cancer: A phase III randomized trial of the Gynecologic Oncology Group. *J. Clin. Oncol.* **31** (suppl.), abstr. 3 (2013).
86. Kindler, H. L. *et al.* Phase II trial of bevacizumab plus gemcitabine in patients with advanced pancreatic cancer. *J. Clin. Oncol.* **23**, 8033–8040 (2005).
87. Ebos, J. M. *et al.* Accelerated metastasis after short-term treatment with a potent inhibitor of tumor angiogenesis. *Cancer Cell* **15**, 232–239 (2009).
88. Páez-Ribes, M. *et al.* Antiangiogenic therapy elicits malignant progression of tumors to increased local invasion and distant metastasis. *Cancer Cell* **15**, 220–231 (2009).
89. Singh, M. *et al.* Anti-VEGF antibody therapy does not promote metastasis in genetically engineered mouse tumour models. *J. Pathol.* **227**, 417–430 (2012).
90. Miles, D. *et al.* Disease course patterns after discontinuation of bevacizumab: pooled analysis of randomized phase III trials. *J. Clin. Oncol.* **29**, 83–88 (2011).
91. Blagoev, K. B. *et al.* Sunitinib does not accelerate tumor growth in patients with metastatic renal cell carcinoma. *Cell Rep.* **3**, 277–281 (2013).
92. Chung, A. S. *et al.* Differential drug class-specific metastatic effects following treatment with a panel of angiogenesis inhibitors. *J. Pathol.* **227**, 404–416 (2012).
93. de Groot, J. F. *et al.* Tumor invasion after treatment of glioblastoma with bevacizumab: radiographic and pathologic correlation in humans and mice. *Neuro Oncol.* **12**, 233–242 (2010).
94. Lu, K. V. *et al.* VEGF inhibits tumor cell invasion and mesenchymal transition through a MET/VEGFR2 complex. *Cancer Cell* **22**, 21–35 (2012).
95. Mellman, I., Coukos, G. & Dranoff, G. Cancer immunotherapy comes of age. *Nature* **480**, 480–489 (2011).
96. Kantoff, P. W. *et al.* Sipuleucel-T immunotherapy for castration-resistant prostate cancer. *N. Engl. J. Med.* **363**, 411–422 (2010).
97. Hodi, F. S. *et al.* Improved survival with ipilimumab in patients with metastatic melanoma. *N. Engl. J. Med.* **363**, 711–723 (2010).
98. Robert, C. *et al.* Ipilimumab plus dacarbazine for previously untreated metastatic melanoma. *N. Engl. J. Med.* **364**, 2517–2526 (2011).
99. Bargou, R. *et al.* Tumor regression in cancer patients by very low doses of a T-cell-engaging antibody. *Science* **321**, 974–977 (2008).
100. Brischwein, K. *et al.* Strictly target cell-dependent activation of T cells by bispecific single-chain antibody constructs of the BiTE class. *J. Immunother.* **30**, 798–807 (2007).
101. Beatty, G. L. *et al.* CD40 agonists alter tumor stroma and show efficacy against pancreatic carcinoma in mice and humans. *Science* **331**, 1612–1616 (2011).
102. Lake, R. A. & Robinson, B. W. S. Immunotherapy and chemotherapy — a practical partnership. *Nature Rev. Cancer* **5**, 397–405 (2005).
103. Zitvogel, L., Apetoh, L., Ghiringhelli, F. & Kroemer, G. Immunological aspects of cancer chemotherapy. *Nature Rev. Immunol.* **8**, 59–73 (2008).
104. Ciampicotti, M., Hau, C.-S., Doornebal, C. W., Jonkers, J. & de Visser, K. E. Chemotherapy response of spontaneous mammary tumors is independent of the adaptive immune system. *Nature Med.* **18**, 344–346 (2012).
105. Zitvogel, L. & Kroemer, G. Reply to: Chemotherapy response of spontaneous mammary tumors is independent of the adaptive immune system. *Nature Med.* **18**, 346 (2012).  
**References 104 and 105 describe contradictory requirements for adaptive immunity in mediating chemotherapeutic responses in preclinical models of cancer.**
106. Arlen, P. M. *et al.* A randomized phase II study of concurrent docetaxel plus vaccine versus vaccine alone in metastatic androgen-independent prostate cancer. *Clin. Cancer Res.* **12**, 1260–1269 (2006).
107. Hamzah, J. *et al.* Vascular normalization in *Rgs5*-deficient tumours promotes immune destruction. *Nature* **453**, 410–414 (2008).
108. Johansson, A., Hamzah, J. J., Payne, C. J. C. & Ganss, R. R. Tumor-targeted TNF $\alpha$  stabilizes tumor vessels and enhances active immunotherapy. *Proc. Natl Acad. Sci. USA* **109**, 7841–7846 (2012).
109. Shrimali, R. K. *et al.* Antiangiogenic agents can increase lymphocyte infiltration into tumor and enhance the effectiveness of adoptive immunotherapy of cancer. *Cancer Res.* **70**, 6171–6180 (2010).
110. Gabrilovich, D. I., Ishida, T., Nadaf, S., Ohm, J. E. & Carbone, D. P. Antibodies to vascular endothelial growth factor enhance the efficacy of cancer immunotherapy by improving endogenous dendritic cell function. *Clin. Cancer Res.* **5**, 2963–2970 (1999).
111. DeSilva, D. R. *et al.* Inhibition of mitogen-activated protein kinase blocks T cell proliferation but does not induce or prevent anergy. *J. Immunol.* **160**, 4175–4181 (1998).
112. Boni, A. *et al.* Selective BRAF<sup>V600E</sup> inhibition enhances T-cell recognition of melanoma without affecting lymphocyte function. *Cancer Res.* **70**, 5213–5219 (2010).
113. Singh, M. & Ferrara, N. Modeling and predicting clinical efficacy for drugs targeting the tumor milieu. *Nature Biotechnol.* **30**, 648–657 (2012).
114. Yang, S. X. *et al.* Gene expression profile and angiogenic marker correlates with response to neoadjuvant bevacizumab followed by bevacizumab plus chemotherapy in breast cancer. *Clin. Cancer Res.* **14**, 5893–5899 (2008).
115. Lambrechts, D., Lenz, H.-J., de Haas, S., Carmeliet, P. & Scherer, S. J. Markers of response for the antiangiogenic agent bevacizumab. *J. Clin. Oncol.* **31**, 1219–1230 (2013).
116. Jubb, A. M. *et al.* Vascular phenotypes in primary non-small cell lung carcinomas and matched brain metastases. *Br. J. Cancer* **104**, 1877–1881 (2011).
117. Sequist, L. V. *et al.* Genotypic and histological evolution of lung cancers acquiring resistance to EGFR inhibitors. *Sci. Transl. Med.* **3**, 75ra26 (2011).
118. Bannoun, J. *et al.* Continuation of bevacizumab after first progression in metastatic colorectal cancer (ML18147): a randomised phase 3 trial. *Lancet Oncol.* **14**, 29–37 (2013).
119. Cheever, M. A. *et al.* Translational Research Working Group developmental pathway for immune response modifiers. *Clin. Cancer Res.* **14**, 5692–5699 (2008).
120. Deng, G., Lu, Y., Zlotnikov, G., Thor, A. D. & Smith, H. S. Loss of heterozygosity in normal tissue adjacent to breast carcinomas. *Science* **274**, 2057–2059 (1996).
121. Patocs, A. *et al.* Breast-cancer stromal cells with TP53 mutations and nodal metastases. *N. Engl. J. Med.* **357**, 2543–2551 (2007).
122. Qiu, W. *et al.* No evidence of clonal somatic genetic alterations in cancer-associated fibroblasts from human breast and ovarian carcinomas. *Nature Genet.* **40**, 650–655 (2008).
123. Giantonio, B. J. *et al.* Bevacizumab in combination with oxaliplatin, fluorouracil, and leucovorin (FOLFOX4) for previously treated metastatic colorectal cancer: results from the Eastern Cooperative Oncology Group Study E3200. *J. Clin. Oncol.* **25**, 1539–1544 (2007).
124. Fiegl, H. *et al.* Breast cancer DNA methylation profiles in cancer cells and tumor stroma: association with HER-2/neu status in primary breast cancer. *Cancer Res.* **66**, 29–33 (2006).
125. Hu, M. *et al.* Distinct epigenetic changes in the stromal cells of breast cancers. *Nature Genet.* **37**, 899–905 (2005).

**Acknowledgements** We would like to express our gratitude to P. Bishop, T. Junttila, K. Leong, J. Settleman, W. Ye, J. Low and S. Scales for their critical review of the manuscript. We also thank the reviewers for valuable insight and suggestions. Our sincere apologies go to authors whose work we are unable to cite due to space limitations.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare competing financial interests see [go.nature.com/hsy7zs](http://go.nature.com/hsy7zs) for details. Readers are welcome to comment on the online version of this article at [go.nature.com/hsy7zs](http://go.nature.com/hsy7zs). Correspondence should be addressed to F.d.S. ([desauvage.fred@gene.com](mailto:desauvage.fred@gene.com)).

# Tumour heterogeneity in the clinic

Philippe L. Bedard<sup>1,2</sup>, Aaron R. Hansen<sup>1,2</sup>, Mark J. Ratain<sup>3</sup> & Lillian L. Siu<sup>1,2</sup>

**Recent therapeutic advances in oncology have been driven by the identification of tumour genotype variations between patients, called interpatient heterogeneity, that predict the response of patients to targeted treatments. Subpopulations of cancer cells with unique genomes in the same patient may exist across different geographical regions of a tumour or evolve over time, called intratumour heterogeneity. Sequencing technologies can be used to characterize intratumour heterogeneity at diagnosis, monitor clonal dynamics during treatment and identify the emergence of clinical resistance during disease progression. Genetic interpatient and intratumour heterogeneity can pose challenges for the design of clinical trials that use these data.**

There is great promise that knowledge of the biological drivers of cancer will lead to personalized cancer treatment. Oncologists increasingly use molecular characterization of a sample of primary or metastatic tumour to guide their selection of treatments for an individual patient. However, they usually rely on a limited sample of cancer tissue that cannot represent heterogeneity between and within patients.

Cancer genomics studies, including large-scale collaborative sequencing projects such as The Cancer Genome Atlas (TCGA) and the International Cancer Genome Consortium (ICGC), have catalogued genetic interpatient tumour heterogeneity for cancers of the same histological subtype. Non-genetic phenotypic and functional heterogeneity is also well recognized (see the Review by Meacham and Morrison on page 328), as is heterogeneity of the tumour microenvironment (see the Review by Junttila and de Sauvage on page 346). Comprehensive characterization of multiple tumour specimens obtained from the same patient illustrates that remarkable intratumour heterogeneity might exist between geographical regions in the same tumour (spatial heterogeneity), as well as between the primary tumour and a subsequent local or distant recurrence in the same patient (temporal heterogeneity). Inter- and intratumour heterogeneity pose a challenge to personalized cancer medicine because a single needle biopsy or surgical excision is unlikely to accurately capture the complete genomic landscape of a patient's cancer. Genomic characterization of cell-free circulating tumour DNA (ctDNA) or circulating tumour cells (CTCs) may offer an opportunity to assess clonal dynamics throughout the course of a patient's illness and identify drivers of therapeutic resistance. Here, we review the clinical implications of interpatient and intratumour heterogeneity for cancer diagnosis, making a prognosis, treatment selection and resistance. We discuss how clinical trials that are restricted to molecular subtypes of cancer could incorporate studies of tumour heterogeneity so that we can better understand the clinical impact of heterogeneity on therapeutic effectiveness and the emergence of treatment resistance.

## Current models for diagnosis and treatment

Modern cancer treatment is based on accurate tissue diagnosis of samples obtained from needle biopsy or surgical excision. Cancerous tissues are analysed under a light microscope to evaluate histopathology, and immunostaining and selected molecular tests are used to establish a specific cancer diagnosis. Treatment is based on the

anatomical location and tissue of origin of the primary tumour when cancer is localized to an organ site or when cancer has metastasized and the primary site can be identified by imaging or pathological examination. When solid tumours recur after treatment for localized disease or progress after systemic treatment for metastatic disease, taking another biopsy to guide treatment decisions is not routine<sup>1</sup>. Instead, further systemic treatment of patients with progressive metastatic disease is typically based on the identification of predictive biomarkers in archived primary specimens, which may no longer represent the current disease such as *BRAF* mutation in melanoma, *HER2* (also known as *ERBB2*) amplification or overexpression in breast cancer, *KRAS* mutation in colorectal cancer and *EGFR* mutation in non-small-cell lung cancer<sup>2–6</sup>.

## Intratumour heterogeneity and clonal evolution

The current approach to molecular biomarker testing to inform cancer treatment focuses on interpatient tumour heterogeneity. However, there is a growing recognition that intratumour heterogeneity within the same patient is clinically relevant because the status of predictive biomarkers that are used for making clinical decisions may evolve during tumour progression, in particular metastatic dissemination of the primary tumour to a distant organ or for established metastatic disease under the selection pressure of treatment. Nowell's theory of clonal evolution states that cancers arise from a single cell of origin, develop genomic instability during replication and then undergo enrichment for the most aggressive clones through the processes of metastasis and the eradication of sensitive clones with cancer treatment<sup>7</sup> (see the Reviews by Burrell *et al.* on page 338 and Klein on page 365). For example, discordance between oestrogen receptor (ER) expression in a primary breast cancer and subsequent distant metastases that may appear many years after completion of primary treatment is observed in 7–25% of patients<sup>8–11</sup> (Table 1). Change in ER status may have important implications for treatment because patients with tumours that lack ER expression do not benefit from treatment with endocrine therapy such as tamoxifen or aromatase inhibitors<sup>12</sup>. Loss of ER or *HER2* expression in breast cancer during metastasis is associated with a poorer outcome<sup>13,14</sup>. Although data from other tumour types are more limited, discordance of prognostic or predictive biomarker testing results between the primary tumour and metastases has been reported in other settings<sup>15–24</sup> (Table 1).

Before metastases are clinically apparent, clonal heterogeneity can

<sup>1</sup>Princess Margaret Cancer Centre, University Health Network, Division of Medical Oncology and Hematology, Toronto, Ontario, Canada. <sup>2</sup>Department of Medicine, University of Toronto, Toronto, Ontario, Canada. <sup>3</sup>Department of Medicine, The University of Chicago, Chicago, Illinois, USA.



**Table 1 | Selected single parameter biomarker tests that are routinely used to inform clinical decision-making for advanced solid tumours, and reported frequencies of discordance between primary tumours and metastases**

Tumour type	Biomarker	Prognostic or predictive biomarker	Evidence of discordance
Oligodendroglioma	1p and 19q co-deletion <i>MGMT</i> promoter methylation	Prognostic/predictive Prognostic/predictive	Not applicable
Medullary thyroid	<i>RET</i> mutation	Prognostic <sup>102</sup>	Unknown
Breast	ER expression	Prognostic/predictive	7–25% <sup>8,11,14</sup>
	PR expression	Prognostic	16–49% <sup>8,11,12,14</sup>
	<i>HER2</i> amplification	Prognostic/predictive	3–24% <sup>13,24</sup>
Lung	<i>EGFR</i> mutation	Prognostic/predictive	0–38% <sup>103,104</sup>
	<i>EML4-ALK</i> translocation	Prognostic/predictive	1–2% <sup>18,105</sup>
Gastric	<i>HER2</i> amplification	Prognostic <sup>106</sup> /predictive <sup>107</sup>	1–3% <sup>20,21</sup>
Colorectal	<i>KRAS</i> mutation	Predictive	0–10% <sup>22,108</sup>
Melanoma	<i>BRAF</i> mutation	Prognostic/predictive	4–25% <sup>109</sup>
Gastrointestinal stromal	<i>KIT</i> mutation	Predictive	Acquired mutations evolve during tyrosine kinase inhibitor treatment <sup>110,111</sup>
	<i>PDGFRA</i> mutation	Predictive	

ER, oestrogen receptor; PR, progesterone receptor

be identified within the primary tumour<sup>25</sup>. For example, complex patterns of *HER2* gene amplification detected by fluorescence *in situ* hybridization are seen in breast<sup>26</sup> and gastro-oesophageal cancers<sup>27</sup>. Similar patterns of regional intratumour heterogeneity have been observed with mutation testing in other tumour types, including *KRAS* in colorectal cancer, *BRAF* in melanoma and *EGFR* in non-small-cell lung cancer<sup>23,28–31</sup>. Intratumour heterogeneity may account for resistance despite the matching of targeted treatment to the mutation, such as trastuzumab for *HER2* amplified breast cancer<sup>32</sup>, *EGFR* monoclonal antibody treatment for *KRAS*-wildtype-colorectal cancer<sup>33</sup> and *EGFR* tyrosine kinase inhibitor treatment for *EGFR* mutant non-small-cell lung cancer<sup>34</sup>, through the selection of clonal subpopulations with mutations that confer treatment resistance.

### Strategies to measure intratumour heterogeneity

Recognition of intratumour heterogeneity to inform treatment decisions requires test methods that can be applied to clinical tumour samples. Genome-scale technologies provide an unbiased characterization of clonal heterogeneity within tumours beyond a specific genetic locus or a set of loci (see the Review by Burrell *et al.* on page 338). Studies with karyotype analysis and comparative genomic hybridization allow for detection of clonal subpopulations within the same tumour that can be differentiated on the basis of DNA content and chromosomal imbalances<sup>35</sup>. Newer techniques such as single nucleotide polymorphism arrays provide greater resolution and can identify smaller-scale allelic imbalances in specific genetic loci. Next-generation sequencing technology allows for the systematic enumeration of single nucleotide mutations and the identification of rare clonal subpopulations that are present in a small fraction of tumour cells. Sequencing studies of normal tissue, early pre-malignant precursors and malignant lesions derived from the same patient have been performed in secondary acute myeloid leukaemia derived from myelodysplastic syndrome<sup>36</sup> and invasive breast cancer with adjacent pre-invasive neoplasia<sup>37</sup>. Clonal lineage has been reconstructed with the identification of antecedent founding clones in a pre-malignant precursor from which malignant disease evolved with the outgrowth of subclones with additional genomic alterations<sup>37</sup>.

In metastatic disease, recent studies have characterized the emergence of treatment-resistant subclones that were present at a minor frequency in the primary tumour<sup>38–44</sup>. This raises the tantalizing possibility that the model of cancer diagnosis and treatment in the future could involve characterization of subpopulations within the primary tumour<sup>45</sup>, monitoring of clonal dynamics during treatment and eradication of treatment-emergent clones. Clinical sequencing using

less invasive sampling methods such as cytology specimens, CTC analysis and ctDNA would greatly facilitate this approach<sup>42,43,46–52</sup>. A recent study<sup>52</sup> demonstrated that ctDNA detected using targeted gene sequencing for *PIK3CA* and *TP53* mutations was associated with survival in patients with metastatic breast cancer. Levels of ctDNA were more closely correlated with response to treatment than CTCs or levels of the circulating cancer antigen CA15-3 detected in serum<sup>53</sup>. A further study involving serial ctDNA exome sequencing of six patients with advanced solid tumours demonstrated an increased representation of certain mutant alleles with the emergence of treatment resistance<sup>54</sup>.

### Challenges of clinical assessment

Beyond initial proof-of-concept studies, larger clinical efforts are required to evaluate whether in-depth genomic characterization and serial monitoring of clonal dynamics leads to better patient care. The falling cost of next-generation sequencing has made high-coverage DNA sequencing of clinically relevant cancer genes accessible at the point of care<sup>55,56</sup>. Genomic assessment of interpatient and intratumour heterogeneity in the clinical environment<sup>57</sup> has several practical challenges.

Surgical resections of primary tumours or metastatic lesions provide large volumes of tumour tissue that are required for assessment of regional heterogeneity and clonal diversity. Tumour specimens are routinely formalin fixed and paraffin embedded (FFPE) after surgical excision to preserve histology. Although tumour nucleic acids can degrade with formalin fixation and this can limit researchers' ability to perform genome-scale analyses, particularly for RNA sequencing, advances in technology mean that the analysis is becoming more feasible. In addition, deciphering the precise spatial orientation of stored FFPE tumour blocks using the routine clinical annotation that is included in surgical pathology reports to reconstruct intratumour heterogeneity can be difficult. Serial characterization of metastatic lesions through core needle biopsy could be used to identify clonal evolution, but sampling bias may occur because only a limited geographical region of a tumour is analysed. ctDNA is more amenable to serial sampling and presumably represents cancer genomes from multiple metastatic sites. However, ctDNA analysis is in its infancy and is not yet routinely established in the clinical environment. Furthermore, whether there are important mutations that are unique to non-circulating populations of tumour cells is not yet known.

In the United States, clinical laboratories that test human specimens for the purpose of providing information on diagnosis, prevention or treatment of disease to the supervising physician must adhere to Clinical Laboratory Improvement Amendments (CLIA) standards and be accredited by the College of American Pathologists (CAP) for reimbursement<sup>58</sup>, and similar regulatory standards exist in

other countries. Genome-scale sequencing was previously outside the purview of a clinical laboratory owing to the cost of massively parallel sequencing platforms, high-performance computing capacity and the sophisticated bioinformatics expertise that was required for sequence alignment and mutation calling. The recent development of bench-top next-generation sequencing instruments that offer high coverage ( $\geq 250 \times$  read depth) of a large targeted panel of clinically relevant cancer genes is well suited to the work flow of a clinical laboratory<sup>59–61</sup>. The Next Generation Sequencing Standardization of Clinical Testing (Nex-StoCT) workgroup recommends that all clinically actionable mutations should be confirmed by independent analysis using an alternative method before reporting to the treating clinician<sup>62</sup>. This poses a problem when high-coverage next-generation sequencing identifies a low-frequency mutation that cannot be confirmed by Sanger or PCR sequencing owing to the limitations of sensitivity of direct sequencing methods.

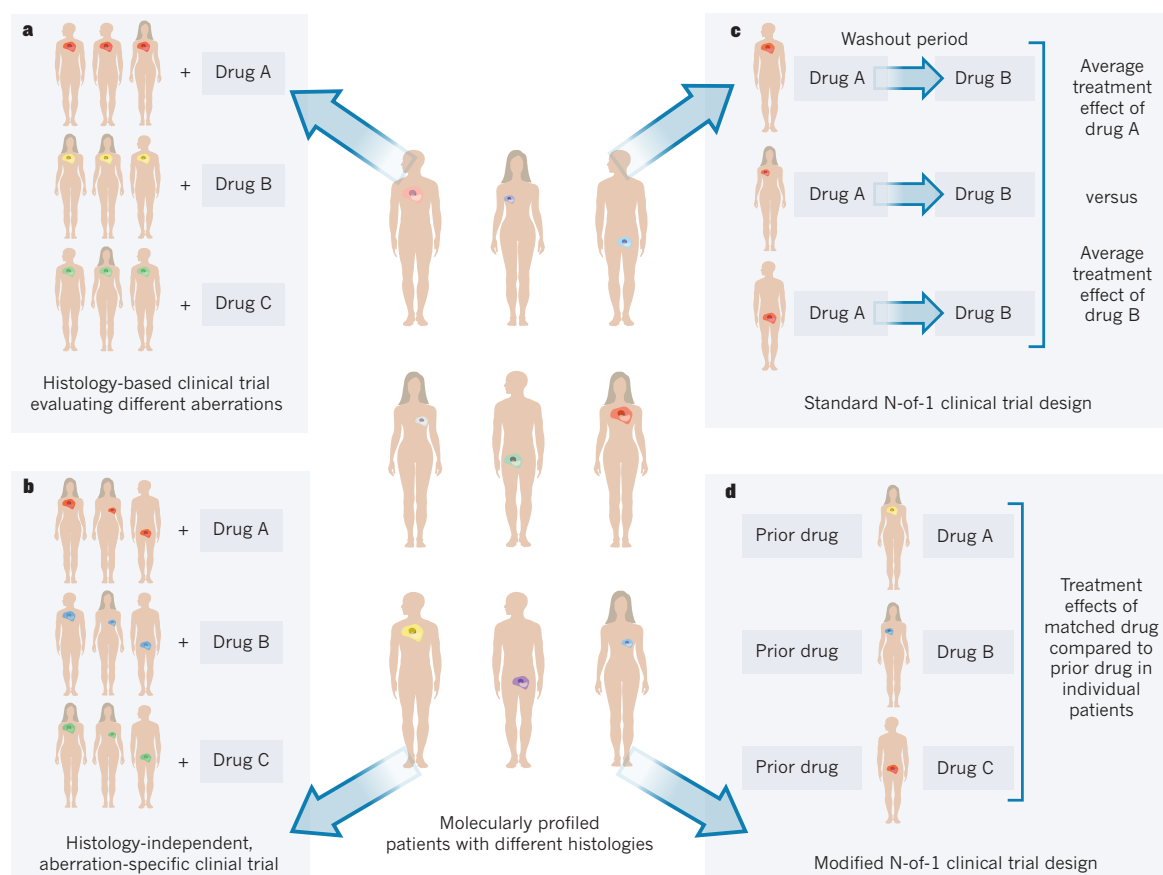
Mutation verification can delay the reporting of results to the oncologist if multiple clinically actionable variants are detected by next-generation sequencing. Patients with metastatic cancer and their oncologists may not be willing to wait for these results before initiating a new treatment<sup>63</sup>. Deciding which mutation or mutations are clinically relevant, prioritizing mutations for treatment matching when multiple mutations are detected and developing a framework to report results to clinicians that can be easily interpreted are complex tasks. Few mutations have been validated with a high level of evidence for the prediction of treatment response<sup>64</sup>. Specific mutations may

have different clinical implications depending on a cancer's tissue of origin, such as *BRAF*(V600E) mutation in patients with melanoma or colorectal cancer<sup>65,66</sup> and their response to vemurafenib monotherapy. For mutations in tumour-specific contexts for which there are no clinical studies available, preclinical drug sensitivity encyclopaedias can be mined to infer potential clinical relevance<sup>67,68</sup>. However, there are concerns about validating predictive genomic biomarkers across cell-line screening data sets<sup>69</sup> and the lack of reproducibility of pre-clinical experiments<sup>70</sup>.

### Trial designs that assess tumour heterogeneity

Despite the challenges associated with genomic assessment in the clinical environment, molecular characterization — from genotyping to targeted genome sequencing — through the use of stored FFPE samples or serially procured fresh tumour biopsies is increasingly used to complement histopathological diagnosis. Clinical-trial design frameworks for cancer diagnostics and therapeutics must be developed to efficiently and dynamically incorporate such genomic data and assess the value of matching profiled patients to specific interventions or targeted therapies.

There are several premises on which clinical-trial design frameworks in the cancer genome era are based. First, genetic aberrations exist in human malignancies with a subset that are present in different cancer types at variable frequencies. Aberrations with functional relevance that lead to cancer initiation, growth and metastasis are the targets of greatest clinical interest because they could potentially be used for diagnosis, prognosis and predicting response to therapy.



**Figure 1 | Clinical-trial design frameworks.** In a population of molecularly profiled patients who have tumours of different histologies (shown by position of tumour) and molecular aberrations (shown as different colours), the framework for a clinical trial can take a number of forms. **a**, Histology-based clinical trials evaluate different molecular aberrations by enrolling patients with the same tumour histology but who harbour different aberrations, and match groups of patients to different drugs. **b**, Histology-independent, aberration-specific clinical trials, or ‘basket’ trials, enrol patients with different tumour histologies but who

harbour the same or related molecular aberrations, and match drugs to the aberration specific or related groups. **c**, Standard N-of-1 trials randomly assign patients to different drugs in different sequential orders, with washout periods between drugs to minimize crossover effects. At completion, the individual effect of each drug and the average effects of each drug across individuals can be analysed. **d**, Modified N-of-1 trials use each patient as his or her own control and compare the treatment effect of the current matched drug with that of the most recent earlier drug.



**Table 2 | Selected worldwide large-scale clinical molecular profiling programmes by institution or consortium**

Trial or programme name	Platforms or techniques	Genes and mutations	Cancer types	Tumour sample
<b>Cancer Research UK, London</b>				
Stratified Medicine Programme <sup>112</sup>	PCR FISH	9 genes 3 genes	Melanoma, NSCLC, CRC and breast, prostate and ovarian cancer	Archival
<b>Dana-Farber Cancer Institute, Boston, Massachusetts</b>				
PROFILE <sup>113</sup>	Sequenom	OncoMap: 41 genes, 471 mutations	All solid tumours	Archival
<b>Curie Institute, Paris; French National Cancer Institute</b>				
SHIVA (NCT01771458)	Ion Torrent PGM CytoScan HD	AmpliSeq: 46 genes 29 genes	All solid tumours	Fresh biopsy
<b>Gustave Roussy Institute, France (non-paediatric trials)</b>				
MOSCATO <sup>75</sup> (NCT01566019)	aCGH PCR	NA 96 mutations	Solid tumour phase I patients	Fresh Biopsy
SAFIRO1 (NCT01414933)	aCGH PCR	NA 2 genes	Breast cancer	Fresh Biopsy
MSN	PCR FISH	Seqcan: 30 genes 5 genes	Melanoma, SCLC and NSCLC	Fresh Biopsy
<b>Massachusetts General Hospital, Boston</b>				
NS <sup>114</sup>	SNAPshot	14 genes, >50 mutations	NSCLC, CRC, melanoma and breast cancer	Archival
<b>MD Anderson Cancer Center, Houston, Texas</b>				
T9 Program <sup>115</sup>	Sequenom	>40 genes	All solid tumours	Archival
IMPACT <sup>73</sup> (NCT00851032)	PCR FISH	10 genes 1 gene	All solid tumours	Archival
Clearing House protocol <sup>116</sup>	PCR Illumina NS, Ion Torrent NS NS	~100 genes T200: 200 genes Whole genome	All solid tumours	Archival or fresh biopsy
<b>Memorial Sloan-Kettering Cancer Center, New York</b>				
IMPACT (NCT01775072)	Illumina HiSeq Sequenom or MiSeq	275 genes (Research assays) NS (Clinical assays)	All solid tumours	Archival
<b>Netherlands</b>				
Centre for Personalized Cancer Treatment <sup>117</sup>	Ion Torrent PGM 5500xl SOLiD	~150 genes >2,000 genes	Solid tumours	Fresh biopsy
<b>Norwegian Cancer Genomics Consortium</b>				
Nationwide programme <sup>118</sup>	NS	Whole exome	9 tumour types, both solid and haematopoietic	Archival or fresh biopsy
<b>Princess Margaret Cancer Centre, Toronto, Canada</b>				
IMPACT <sup>60</sup> (NCT01505400)	MiSeq Sequenom	TSACP: 48 genes, >700 mutations. Customized panel: 23 genes, 279 mutations	Selected solid tumours	Archival
<b>Vall d'Hebron Institute of Oncology, Barcelona, Spain</b>				
NS <sup>72, 119</sup>	Sequenom  Illumina GAIx	OncoCarta, 19 genes, 238 mutations NS	Breast cancer, solid tumour phase I patients	Archival
<b>Vanderbilt-Ingram Cancer Center, Nashville, Tennessee</b>				
PCMI <sup>120</sup>	SNAPshot	6–8 genes and >40 mutations	Melanoma, NSCLC, CRC and breast cancer	Archival
<b>WIN Consortium</b>				
WINTHER <sup>83</sup> (NCT01856296)	NGS CNV CGH	NS NS NA	Solid tumours	Fresh biopsy (tumour and matched normal)

aCGH, array comparative genomic hybridization; CGH, comparative genomic hybridization; CNV, copy number variation; CRC, colorectal cancer; FISH, fluorescence *in-situ* hybridization; GAIx, genome analyzer Ix; NA, not applicable; NGS, next-generation sequencing; NS, not stipulated; NSCLC, non-small-cell lung cancer; PCR, polymerase chain reaction; PCMI, personalized cancer medicine initiative; PGM, personal genome machine; SCLC, small-cell lung cancer; TSACP, TruSeq amplicon Cancer Panel.

Second, there are specific interventions or tolerable medicinal agents that may effectively modulate such targets. Last, intratumour heterogeneity and clonal evolution occur and there are feasible technologies to measure these phenomena in the clinical setting. The reliable quantification of both spatial and temporal variations in

the molecular landscape of cancers would enable the development of therapeutic strategies to interrogate them. Although, currently, most approved targeted therapies and clinical trials focus on inter-patient heterogeneity, considering intratumour heterogeneity will increasingly become important in the future.

## Trial designs for interpatient heterogeneity

Clinical-trial design frameworks that focus on interpatient tumour heterogeneity are possible, assuming that detailed genomic characterization is feasible<sup>71</sup> (Fig. 1).

### Longitudinal cohort with nested trials

One framework currently used by many large cancer institutions and national cancer cooperative groups is to prospectively profile a large number of patients to establish a longitudinal cohort with clinical annotation such as demographics, histopathological diagnosis, earlier therapies and outcome (Table 2). Thus far, most clinical molecular profiling programmes worldwide have focused on the genomic characterization of limited but presumably representative specimens obtained at a single time point, typically in patients with metastatic disease who are suitable for systemic therapy. It is logical that a current sample would more accurately reflect the patient's current disease than an archived sample, although it is unknown whether a small current specimen (for example, from a needle biopsy) is preferable to a larger historical sample (for example, from surgical resection). In instances in which archived tumour tissue is too scant to yield sufficient DNA, or has been exhausted owing to serial evaluations of single markers, then a fresh tumour biopsy would be necessary for genomic profiling. Although there is great enthusiasm for molecular characterization of tumour samples, the clinical use of this approach is still unproven. Some clinical trials of targeted drugs limit enrolment to patients with specific molecular perturbations; however, the effectiveness of such drugs is usually unconfirmed. Nonetheless, the coupling of a molecular characterization strategy with a drug development programme has been widely embraced, despite disparate results from different retrospective series and the lack of definitive supportive data<sup>72,73</sup>. In this context, patients with specific molecular aberrations are often 'opportunistically' enrolled into clinical trials of matching targeted agents. This framework is attractive to those running large programmes who have access to a robust panel of early phase clinical trials that test different molecular targets<sup>60,63,73–75</sup>. The panel of clinical trials can be 'nested' or embedded as distinct research activities under the auspices of an overarching platform of molecular profiling and target–drug matching.

### Histology-based design

Other frameworks involve the evaluation of the target–agent matching strategy in large, prospectively conducted clinical trials (Table 2). For instance, histology-based and biomarker-integrated multicentre clinical trials aim to assess a variety of targeted agents matched to specific molecular profiles within a single tumour type (Fig. 1a). The FOCUS 4 trial supported by the UK Medical Research Council, for example, will enrol patients with advanced colorectal cancer who have responsive or stable disease after 16 weeks of chemotherapy<sup>76</sup>. On molecular profiling, patients with tumours that harbour commonly mutated oncogenes such as *KRAS*, *BRAF* or *PIK3CA* will be given targeted agents or a placebo. Other histology-based clinical trials include the US-based BATTLE-2 trial (NCT01248247) in lung cancer and the I-SPY 2 trial in breast cancer (NCT01042379)<sup>77–79</sup> (Table 3).

### Histology-agnostic, aberration-specific design

An alternative framework employs a histology-agnostic, aberration-specific design in which patients whose tumours harbour identical or related molecular profiles are treated in the same 'basket' with the same therapeutic regimen (Fig. 1b). An example is the inclusion of different tumour types that harbour *PIK3CA* mutations or amplifications into a basket trial that evaluates a PI(3)K  $\alpha$ -isoform specific inhibitor (NCT01219699). This strategy may be adapted to increase enrolment of patients with tumour types that demonstrate early signals of antitumour activity while excluding those who lack preliminary response. Although this framework will not directly lead to regulatory approval, given its exploratory nature, it does provide a platform to determine the differences in functionality of the same molecular alteration across multiple cancer types.

## N-of-1 clinical trial design

The N-of-1 clinical-trial design framework has been pursued for non-oncology diseases, most frequently in neuropsychiatric, pulmonary and musculoskeletal conditions<sup>80,81</sup>. In their standard context, N-of-1 trials involve individual patients who are typically blinded and randomly assigned to different treatment regimens or to a placebo in different sequential orders, with washout periods, in which patients receive no treatment, between regimen alterations to minimize crossover effects (Fig. 1c). There are limitations to the application of this framework in oncology. For instance, the switch from one regimen to another may occur before there is sufficient time for antitumour activity to be manifested, such that there may be an underestimation of therapeutic efficacy while increasing the risk of inducing drug resistance. Modified N-of-1 designs have been used to investigate the value of individualized therapy. The concept of using each individual patient as his or her own control, for example, to assess the growth modulation index by comparing the time to progression or progression-free survival (PFS) on a current regimen with that attained on the most recent prior treatment, represents such a modification of the N-of-1 design<sup>82</sup> (Fig. 1d). This framework may become increasingly relevant for subsets of patients with rare molecular alterations, for which large randomized trials may never be feasible. The WINTHER trial<sup>83</sup> (NCT01856296), led by the Worldwide Innovative Networking (WIN) Consortium, is an example of a modified N-of-1 design that is using a variety of advanced profiling technologies to comprehensively characterize oncogenic events in 200 patients with different cancers. The trial compares patients' PFS on therapy guided by profiling results with that achieved on the regimen immediately preceding trial enrolment. However, the validity of this approach is unknown, given the uncertain correlation in PFS between sequential inactive therapies<sup>82</sup>.

## Trial designs for intratumour heterogeneity

Establishing clinical-trial design frameworks in the context of inpatient tumour heterogeneity and clonal evolution is challenging because dimensions of both space and time must be incorporated to reflect the dynamic nature of tumour biological characteristics within individuals.

### Geographical heterogeneity

The execution of the aforementioned frameworks is typically based on molecular profiling of tumour specimens obtained from one geographical location. These samples, in addition to other biorepositories such as tissue banks and autopsy programmes, provide a means to build knowledge bases that help us to gain insight into complex molecular events such as intratumour heterogeneity<sup>40,84,85</sup>. One such initiative to build this type of knowledge base is the REACT study (NCT01505400). The aim of this study is to genomically evaluate all archived tumour samples from a cohort of molecularly profiled patients to assess heterogeneity and clonal evolution.

To prospectively assess geographical or spatial heterogeneity, profiling of multiregional tumour samples would be indicated. Although this is feasible (but rarely performed in surgical resections), it is impractical and potentially risky to take biopsies from multiple deep-seated metastatic lesions in every patient to examine the genotypes of different tumour cell clones. If tumour biopsies using fine-gauge needles (23-gauge or smaller) could yield sufficient quantities of tumour nucleic acids for molecular profiling, these would be an attractive alternative to large-bore needles owing to the lower risk of procedure-related complications. An ongoing prospective study called MATCH (NCT01703585) evaluating the quality and quantity of DNA obtained using different sizes of biopsy needles, could determine whether the use of fine-gauge needles is feasible for targeted sequencing. In addition, there are prospective and retrospective tumour-specific programmes that explore heterogeneity and evolution in relation to drug therapy (Table 3). The PREDICT programme for patients with renal cell cancer who are treated with neoadjuvant everolimus or sunitinib is an example of an explicitly designed study to evaluate heterogeneity in the primary tumour



**Table 3 | Prospective and retrospective programmes that evaluate tumour heterogeneity by institution or consortium**

Trial or programme	Platforms or techniques	Coverage and depth	Cancer types	Fresh tumour acquisition	Additional specimen collection	Drugs	Additional procedures	Type of heterogeneity
<b>Dana-Farber Cancer Institute, Boston, Massachusetts; Broad Institute, Cambridge, Massachusetts; and Brigham and Women's Hospital, Boston</b>								
CanSeq <sup>121</sup>	Sequenom NS	OncoMap Whole exome, depth NS	NSCLC, CRC, MBC and prostate cancer	Biopsy at set time point depending on tumour type	No	No	No	Interpatient
<b>Heinrich Heine University, Düsseldorf</b>								
DETECT III <sup>101</sup> (NCT01619111)	NS	NS	HER-2 <sup>+</sup> MBC with HER2 <sup>+</sup> CTCs	No	CTCs (CellSearch)	SOC chemotherapy or endocrine therapy +/- lapatinib	No	Intratumour
<b>Massachusetts General Hospital, Boston</b>								
Biopsies of Cancer Patients for Tumor Molecular Characterization (NCT01061944)	NS	Genes NS, depth NS	All solid tumours	Biopsy of metastasis (SOC)	NS	No	No	Interpatient
<b>Mayo Clinic, Scottsdale, Arizona</b>								
BEAUTY <sup>122</sup>	NS	Whole genome, depth NS	Non-metastatic breast cancer	Biopsy of primary pre- and post-neoadjuvant chemotherapy. Primary resection	NS	Paclitaxel +/- trastuzumab AC or FEC	Xenografts	Intratumour Interpatient
<b>MRC Clinical Trials Unit, London</b>								
FOCUS 4 (ref. 76)	PCR assays	3 genes	CRC	Diagnostic, on treatment and PD biopsy (lesion NS)	No	5 treatment arms	No	Interpatient
<b>MD Anderson Cancer Center, Houston, Texas</b>								
BATTLE-2 (NCT01248247)	PCR FISH	11 biomarkers	NSCLC, PD on chemotherapy	Biopsy (lesion NS)	NS	Erlotinib, MK2206, AZD6224, sorafenib	No	Interpatient
BATTLE-Front line (NCT01263782)			NSCLC, treatment naive					Interpatient
<b>National Cancer Institute, Bethesda, MD</b>								
I-SPY 2 (NCT01042379)	TargetPrint HER2, Mamma-print	71 genes	Stage 3 breast cancer	Pre-specified serial primary biopsies. Primary resection	Blood	Experimental drugs with SOC chemotherapy	Breast MRI	Interpatient
MPACT <sup>123</sup>	NS	22 genes for treatment, 80kb sequenced, 383 amplicons with ≥80% covered >450x	All solid tumours	Biopsy of metastasis	NS	MEK, mTOR, PARP, WEE1 inhibitors	No	Interpatient
MATCH <sup>124</sup>	NS	Genes and depth NS	All solid tumours and lymphoma. PD on 1 SOC treatment	Pretreatment and PD biopsy (lesion NS)	NS	Multiple targeted therapies on clinical trials	RNA-Seq	Intratumour Interpatient
<b>PREDICT Consortium</b>								
E-PREDICT S-PREDICT (ISRCTN 22979604)	GAllx HiSeq	Whole exome, transcriptome, average depth 30x	Renal cell cancer	Biopsy of primary and metastasis. Nephrectomy	NS	Everolimus or sunitinib	Functional RNA interference	Intratumour Interpatient
<b>Princess Margaret Cancer Centre, Toronto, Canada</b>								
MATCH (NCT01703585)	MiSeq	48 genes, 212 amplicons average depth ~1,000x	CRC, MBC, gynaecological cancers	Serial biopsies of metastases at study start and on PD	Blood: CTCs, ctDNA. Archival tumour	No	No	Intratumour Interpatient
	Sequenom	23 genes, 279 hotspots						
<b>University College London</b>								
TRACERx (NCT01888601)	NS	Whole genome, whole exome, depth NS	NSCLC	Biopsy of primary and metastasis. Primary resection	Blood: ctDNA	SOC chemotherapy	Functional imaging	Intratumour Interpatient

AC, adriamycin and cyclophosphamide; CRC, colorectal cancer; CTCs, circulating tumour cells; ctDNA, circulating tumour DNA; FEC, 5-fluorouracil, epirubicin, cyclophosphamide; FISH, fluorescence *in situ* hybridization; GAllx, genome analyzer IIx; kb, kilobases; MBC, metastatic breast cancer; MRI, magnetic resonance imaging; NS, not stipulated; NSCLC, non-small-cell lung cancer; PCR, polymerase chain reaction; PD, progressive disease; SOC, standard of care.

**Table 4 | The different clinical-trial design frameworks and tumour-sampling strategies that can be used to evaluate intratumour heterogeneity and clonal evolution from pre-malignancy to the development of resistant metastases, using breast cancer as an example.**

Ductal carcinoma <i>in situ</i>	Localized cancer	Micrometastases	Macrometastases	Resistant disease
<b>Clinical evaluation</b>				
Surveillance of pre-malignancy to malignancy	Forming a prognosis of metastatic potential	Monitoring response to adjuvant therapy	Targeting treatment to match driver clones	Targeting treatment to match resistant clones
<b>Evaluation strategies</b>				
Multiregional sampling, if feasible	Multiregional sampling Monitoring using CTCs or ctDNA	Monitoring using CTCs or ctDNA	Multiregional sampling Molecular imaging Serial sampling CTCs or ctDNA	Multiregional sampling Serial sampling for CTCs or ctDNA Molecular imaging
<b>Clinical-trial design frameworks</b>				
Longitudinal cohort	Longitudinal cohort Histology-based design	Longitudinal cohort Histology-based design	Histology-based design Histology-agnostic basket design	Histology-based design Histology-agnostic basket design N-of-1 design

CTC, circulating tumour cell; ctDNA, circulating tumour DNA.

through multiregional sampling<sup>86,87</sup>. Ultimately, the development of non-invasive visualization techniques, such as molecular imaging using radionuclide-based methods that can quantify the expression of tumoral targets with high sensitivity and specificity, would be ideal<sup>88</sup>.

### Temporal heterogeneity

Serial tumour sampling, especially at crucial time points in the disease course such as the development of metastatic disease or progression after initial response to systemic therapy, may reveal the emergence of dominant clones. This type of dynamic examination of clonal evolution is being conducted by programmes such as PREDICT<sup>86,87</sup>. Until less invasive techniques such as characterization of CTCs or ctDNA are validated to yield sufficient sensitivity and specificity to be representative of clonal distribution and evolutionary pattern, fresh tumour biopsies will probably be used to monitor these events, although limited biopsies may also not reflect the full genomic landscape<sup>89,90</sup>. Sensitivity of detection of somatic mutations is related to their frequencies in the analysed segments of cancer-related genes, and can be increased by using new techniques such as amplification and deep sequencing of selected genomic regions. ctDNA has already been used as a tracking tool for distinct existent clones<sup>51</sup>, as well as an early predictor of treatment response or resistance<sup>42,43,52</sup>. Optimization of these methods to transition them from research to diagnostic laboratories would enable their applications in clinical trials and eventually in routine cancer care. Advances in molecular imaging that would make longitudinal surveillance possible would be desirable, although it is uncertain whether imaging can ever provide resolution at the level of target expression in tumour cells and be able to reflect changes in the clonal milieu.

### Trial designs

The evaluation of geographical and temporal variations in tumour molecular profiles is complex. Their integration into the aforementioned clinical-trial design frameworks that focus on interpatient tumour heterogeneity is possible but would necessitate that the frameworks become dynamic models that consider changes across space and time within individuals. By using these frameworks, geographical heterogeneity and clonal evolution in tumour samples can be prospectively measured, but must first be correlated with clinical outcome to determine whether they portend a prognostic and/or predictive role. Interventions to modulate these phenomena would only be planned if they are demonstrated to have an important link to clinical outcome.

### An example of intratumour heterogeneity in the clinic

Given the complexity of intratumour heterogeneity and clonal evolution, it is impossible to provide approaches that are universally applicable. As such, in reality, it is expected that adaptations of clinical trial designs for individuals will be tailored to the unique features of specific malignancies. By using breast cancer as a example, practical

applications of these frameworks for prognosis and therapy (Table 4) are discussed below.

### Pre-malignancy to malignancy

The establishment of a longitudinal cohort would enable long-term follow-up of patients with pre-malignant lesions, such as ductal carcinoma *in situ* (DCIS), for whom the disease might progress to invasive breast cancer. Retrospective analyses of cases with synchronous DCIS and invasive ductal carcinoma have shown that this progression is associated with the appearance of subclones that harbour specific genetic aberrations, such as amplifications of *MYC*, *CCND1* and *FGFR1* (refs 91–93). The prospective quantification of geographical and temporal heterogeneity can be achieved by multiregional sampling of DCIS in surgical specimens, and by serial sampling in cases of DCIS recurrence. The identification of biomarkers of progression that may predict the transition from pre-malignancy to malignancy would be relevant. A comparison of surveillance strategies with or without molecular assessment of such biomarkers in different geographical locations and in serially collected samples of pre-malignant lesions can be undertaken to validate their prognostic role.

### Metastatic potential of localized cancer

Both the longitudinal cohort strategy and the histology-based design to evaluate multiple aberrations would be reasonable frameworks to consider for metastatic potential of localized cancer. Comprehensive molecular portraits of the four main primary breast cancer subtypes (luminal A, luminal B, basal-like and HER2-enriched) have recently been published<sup>94</sup>. Multiregional sampling and molecular profiling of primary tumour and regional lymph nodes can be carried out in patients who have undergone curative resections. In addition, depending on the sensitivity of detection, CTCs can be enumerated and profiled, and ctDNA can be extracted and analysed for the presence of somatic genomic alterations. Patients can then be monitored prospectively to determine if the detection of specific biomarkers in multiple locations within the primary surgical specimen or in the circulatory system can help to identify those tumours with biologically aggressive behaviour beyond the prognosis given by standard clinicopathological factors.

### Monitoring for early micrometastases

After definitive local therapy and systemic adjuvant therapy, serial enumeration of CTCs or prospective sequential profiling of ctDNA can be performed, either as a longitudinal cohort or in a histology-based design to evaluate different molecular aberrations<sup>51</sup>. Single-cell exome sequencing to detect single nucleotide mutations is being developed<sup>95–97</sup>, such that molecular characterization using captured CTCs could eventually be possible<sup>98</sup>. These samples can be used as a tracking tool for distinct existent clones that can be assessed to monitor response to adjuvant therapy and to predict disease relapse.



## Targeting oncogenic driver clones

In patients who develop macroscopic metastases from breast cancer, current systemic therapy consists mainly of hormonal therapy, cytotoxic chemotherapy and a limited number of targeted agents, such as HER2 inhibitors for HER2-positive tumours, or mTOR inhibitors in hormone-receptor-positive tumours<sup>99</sup>. At present, other than HER2-targeting, selecting treatment based on a molecular profile is not proven to be superior to standard algorithms in metastatic breast cancer. As such, the design of therapeutic clinical trials that are either histology-based or histology-agnostic to evaluate the benefit of target–drug matching compared with conventional approaches, would be considered investigational. Exploring the impact of intratumour heterogeneity in a therapeutic context adds a further layer of complexity. Even if current technologies such as minimally invasive multiregional sampling of metastases or molecular imaging are able to identify functional tumour subpopulations that are geographically distinct, the design of clinical trials to interrogate these subpopulations is challenging. For instance, if two potentially important clones, one with *PIK3CA* mutations and the other with *FGFR1* amplification coexist, then hypothetical therapeutic possibilities can include either concurrent combination or sequential treatment with PI(3)K and FGFR1 inhibitors (ideally distinguished using carefully designed randomized trials). The accessibility to approved or experimental agents in such scenarios may be limited. Furthermore, the most optimal approach to combine or sequence two or more agents to yield sufficient biological target modulation with tolerable toxicity is often undefined and requires dose-finding studies. Finally, even if appropriate drug combination strategies are determined and can effectively suppress clonal evolution, thus ameliorating or delaying the onset of resistance, a previously undetected or new driver clone may ultimately arise. In contrast to the uncertainty of ‘drugging’ intratumour heterogeneity successfully, the use of CTCs or ctDNA as early biomarkers of treatment response of metastatic breast cancer seems to be more readily tangible<sup>52</sup>.

## Emergence of resistant clones

Intratumour heterogeneity is a key factor that may lead to primary drug resistance because the extent of genomic assessment and molecular characterization determines our ability to identify potentially important subclones<sup>100</sup>. In patients who have clearly responded to treatment but in whom disease subsequently progresses, a repeat tumour biopsy to detect the expansion of pre-existent resistant subclones or the emergence of newly acquired resistant clones may be highly informative. An important caveat is that clonal population size and architecture cannot be assessed through biopsy sampling of a single metastatic site. If a change in genotype is observed when another biopsy is taken at the onset of progression after systemic treatment, this may be due to either clonal evolution or as a result of an earlier false negative due to sampling bias. To circumvent such limitations of tumour biopsies, characterization of CTC or ctDNA in plasma could be an attractive alternative if they are demonstrated to be more reflective of the global molecular status. Furthermore, these circulating ‘liquid tumours’ may also precede radiological evidence of tumour growth<sup>42,43,54</sup>. These strategies to identify and tackle primary or acquired resistance can be integrated into clinical trials using histology-based or histology-agnostic frameworks. For instance, DETECT III (NCT01619111)<sup>101</sup> is a multicentre, histology-based, randomized phase III study that compares lapatinib (as a HER2-targeted therapy) combined with standard therapy with standard therapy alone in patients with HER2-negative breast cancer who have had HER2-positive CTCs detected in their blood. When the sample size is small, an N-of-1 trial design may be used to sequentially assess, in the same patient, the effects of different agents that may have antitumour activity against the resistant clones. It would seem logical to interrogate an emerging resistant clone as early as possible, using the combination or sequential therapeutic strategies previously described, although the timing for pharmacological counteraction of clonal evolution may also require full assessment through well-conceived clinical trials.

## Future directions

The occurrence of intratumour heterogeneity and clonal evolution in cancers, resulting in malignant growth, invasion, metastasis and resistance acquisition has long been recognized. The availability of molecular profiling technologies such as next-generation sequencing coupled with advances in bioinformatics has enabled these previously elusive phenomena to be assayed in the clinical setting. The challenges ahead are immense, and include the reliable and accurate elucidation of geographical and temporal variations in patient samples and the subsequent correlation with both prognosis and treatment response. Current efforts are focused on gathering evidence to support the idea that intratumour heterogeneity substantially affects disease outcome, although the relationship is probably context dependent. Clinical trial strategies to interrogate intratumour heterogeneity are challenging, and for researchers to gain a deeper understanding into these molecular complexities would require not only the active participation of patients who are willing to undergo repeated investigations, but also the collaborative engagement of clinicians and scientists. Without a full understanding of the spectrum of a patient’s mutations, we may risk expending large resources on the development of fundamentally flawed approaches to biomarker-directed therapeutics.

The knowledge that significant intratumour heterogeneity is present in most patients has important implications for predictive biomarker development in the context of early clinical trials. First, quantitative biomarkers (for example, RNA expression) may be misleading, as they are based on the average expression across a heterogeneous tumour. Second, sequencing approaches may be misleading, unless careful attention is paid to detecting minor clones of clinical significance. Last, phenotypic and functional heterogeneity that results from events other than genomic alterations, for instance due to epigenetic alterations or plasticity, is likely to have an important effect on treatment response (see Review by Meacham and Morrison on page 328). Although we do not yet have the knowledge base to successfully individualize treatment by accounting for both interpatient and inpatient heterogeneity, we believe that the delivery of comprehensive personalized cancer medicine will eventually be possible. ■

Received 9 April; accepted 25 July 2013.

1. Tran, B. *et al.* Cancer genomics: technology, discovery, and translation. *J. Clin. Oncol.* **30**, 647–660 (2012).
2. Mok, T. S. *et al.* Gefitinib or carboplatin–paclitaxel in pulmonary adenocarcinoma. *N. Engl. J. Med.* **361**, 947–957 (2009).
3. De Roock, W. *et al.* KRAS wild-type state predicts survival and is associated to early radiological response in metastatic colorectal cancer treated with cetuximab. *Ann. Oncol.* **19**, 508–515 (2008).
4. Chapman, P. B. *et al.* Improved survival with vemurafenib in melanoma with BRAF V600E mutation. *N. Engl. J. Med.* **364**, 2507–2516 (2011).
5. Slamon, D. J. *et al.* Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *N. Engl. J. Med.* **344**, 783–792 (2001).
6. Karapetis, C. S. *et al.* K-ras mutations and benefit from cetuximab in advanced colorectal cancer. *N. Engl. J. Med.* **359**, 1757–1765 (2008).
7. Nowell, P. C. The clonal evolution of tumor cell populations. *Science* **194**, 23–28 (1976).
8. Amir, E. *et al.* Prospective study evaluating the impact of tissue confirmation of metastatic disease in patients with breast cancer. *J. Clin. Oncol.* **30**, 587–592 (2012).
9. Gong, Y., Han, E. Y., Guo, M., Pusztai, L. & Sneige, N. Stability of estrogen receptor status in breast carcinoma. *Cancer* **117**, 705–713 (2011).
10. Thompson, A. M. *et al.* Prospective comparison of switches in biomarker status between primary and recurrent breast cancer: the Breast Recurrence In Tissues Study (BRITS). *Breast Cancer Res.* **12**, R92 (2010).
11. Chang, H. J. *et al.* Discordant human epidermal growth factor receptor 2 and hormone receptor status in primary and metastatic breast cancer and response to trastuzumab. *Jpn. J. Clin. Oncol.* **41**, 593–599 (2011).
12. Early Breast Cancer Trialists’ Collaborative Group. Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. *Lancet* **378**, 771–784 (2011).
13. Lindström, L. S. *et al.* Clinically used breast cancer markers such as estrogen receptor, progesterone receptor, and human epidermal growth factor receptor 2 are unstable throughout tumor progression. *J. Clin. Oncol.* **30**, 2601–2608 (2012).
14. Liedtke, C. *et al.* Prognostic impact of discordance between triple-receptor measurements in primary and recurrent breast cancer. *Ann. Oncol.* **20**, 1953–1958 (2009).

15. Tapia, C. *et al.* *HER2* gene status in primary breast cancers and matched distant metastases. *Breast Cancer Res.* **9**, R31 (2007).
16. Artale, S. *et al.* Mutations of *KRAS* and *BRAF* in primary and matched metastatic sites of colorectal cancer. *J. Clin. Oncol.* **26**, 4217–4219 (2008).
17. Kalikaki, A. *et al.* Comparison of *EGFR* and *K-RAS* gene status between primary tumours and corresponding metastases in NSCLC. *Br. J. Cancer* **99**, 923–929 (2008).
18. Kim, H. *et al.* Discordance between anaplastic lymphoma kinase status in primary non-small-cell lung cancers and their corresponding metastases. *Histopathology* **62**, 305–314 (2013).
19. Weller, M. *et al.* *MGMT* promoter methylation in malignant gliomas: ready for personalized medicine? *Nature Rev. Neurol.* **6**, 39–51 (2010).
20. Bozzetti, C. *et al.* Comparison of *HER2* status in primary and paired metastatic sites of gastric carcinoma. *Br. J. Cancer* **104**, 1372–1376 (2011).
21. Kim, M. A., Lee, H. J., Yang, H. K., Bang, Y. J. & Kim, W. H. Heterogeneous amplification of *ERBB2* in primary lesions is responsible for the discordant *ERBB2* status of primary and metastatic lesions in gastric carcinoma. *Histopathology* **59**, 822–831 (2011).
22. Knijn, N. *et al.* *KRAS* mutation analysis: a comparison between primary tumours and matched liver metastases in 305 colorectal cancer patients. *Br. J. Cancer* **104**, 1020–1026 (2011).
23. Yancovitz, M. *et al.* Intra- and inter-tumor heterogeneity of *BRAF*<sup>V600E</sup> mutations in primary and metastatic melanoma. *PLoS ONE* **7**, e29336 (2012).
24. Niikura, N. *et al.* Loss of human epidermal growth factor receptor 2 (*HER2*) expression in metastatic sites of *HER2*-overexpressing primary breast tumors. *J. Clin. Oncol.* **30**, 593–599 (2012).
25. Fidler, I. J. & Hart, I. R. Biological diversity in metastatic neoplasms: origins and implications. *Science* **217**, 998–1003 (1982).
26. Starczynski, J. *et al.* *HER2* gene amplification in breast cancer: a rogues' gallery of challenging diagnostic cases: UKNEQAS interpretation guidelines and research recommendations. *Am. J. Clin. Pathol.* **137**, 595–605 (2012).
27. Yoon, H. H. *et al.* Adverse prognostic impact of intratumor heterogeneous *HER2* gene amplification in patients with esophageal adenocarcinoma. *J. Clin. Oncol.* **30**, 3932–3938 (2012).
28. Perez, K. *et al.* Heterogeneity of colorectal cancer (CRC) in reference to *KRAS* proto-oncogene utilizing WAVE technology. *Exp. Mol. Pathol.* **95**, 74–82 (2013).
29. Wilmott, J. S. *et al.* Intratumoral molecular heterogeneity in a *BRAF*-mutant, *BRAF* inhibitor-resistant melanoma: a case illustrating the challenges for personalized medicine. *Mol. Cancer Ther.* **11**, 2704–2708 (2012).
30. Taniguchi, K., Okami, J., Kodama, K., Higashiyama, M. & Kato, K. Intratumor heterogeneity of epidermal growth factor receptor mutations in lung cancer and its correlation to the response to gefitinib. *Cancer Sci.* **99**, 929–935 (2008).
31. Sakurada, A., Lara-Guerra, H., Liu, N., Shepherd, F. A. & Tsao, M.-S. Tissue heterogeneity of *EGFR* mutation in lung adenocarcinoma. *J. Thorac. Oncol.* **3**, 527–529 (2008).
32. Rye, I. *et al.* Intra-tumor heterogeneity as a predictor of therapy response in *HER2* positive breast cancer. *Cancer Res.* **72**, (24 suppl.) P3–05–04 (2012).
33. Tougeron, D. *et al.* Effect of low-frequency *KRAS* mutations on the response to anti-*EGFR* therapy in metastatic colorectal cancer. *Ann. Oncol.* **24**, 1267–1273 (2013).
34. Bai, H. *et al.* Detection and clinical significance of intratumoral *EGFR* mutational heterogeneity in Chinese patients with advanced non-small cell lung cancer. *PLoS ONE* **8**, e54170 (2013).
35. Coons, S. W., Johnson, P. C. & Shapiro, J. R. Cytogenetic and flow cytometry DNA analysis of regional heterogeneity in a low grade human glioma. *Cancer Res.* **55**, 1569–1577 (1995).
36. Walter, M. J. *et al.* Clonal architecture of secondary acute myeloid leukemia. *N. Engl. J. Med.* **366**, 1090–1098 (2012).
- This study reconstructs the clonal evolution of secondary acute myeloid leukaemia from myelodysplastic syndrome, demonstrating that a single population of myelodysplastic syndrome cells underwent multiple rounds of mutation and selection.**
37. Newburger, D. E. *et al.* Genome evolution during progression to breast cancer. *Genome Res.* **23**, 1097–1108 (2013).
38. Shah, S. P. *et al.* Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. *Nature* **461**, 809–813 (2009).
- This was the first study to demonstrate how genome sequencing can be used to follow the evolution of a solid tumour from early to advanced stage.**
39. Yachida, S. *et al.* Distant metastasis occurs late during the genetic evolution of pancreatic cancer. *Nature* **467**, 1114–1117 (2010).
40. Campbell, P. J. *et al.* The patterns and dynamics of genomic instability in metastatic pancreatic cancer. *Nature* **467**, 1109–1113 (2010).
- References 39 and 40 describe how genomic instability drives clonal evolution in metastatic pancreatic cancers.**
41. Ding, L. *et al.* Genome remodelling in a basal-like breast cancer metastasis and xenograft. *Nature* **464**, 999–1005 (2010).
42. Diaz, L. A. Jr *et al.* The molecular evolution of acquired resistance to targeted *EGFR* blockade in colorectal cancers. *Nature* **486**, 537–540 (2012).
43. Misale, S. *et al.* Emergence of *KRAS* mutations and acquired resistance to anti-*EGFR* therapy in colorectal cancer. *Nature* **486**, 532–536 (2012).
- References 42 and 43 demonstrate how *KRAS*-mutant subclones in colorectal cancer can emerge under the selection pressure of *EGFR* antibody therapy to confer treatment resistance.**
44. Shah, S. P. *et al.* The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *Nature* **486**, 395–399 (2012).
- This study demonstrates wide variety of clonal and mutational composition in 104 primary triple-negative breast cancers.**
45. Aparicio, S. & Caldas, C. The implications of clonal genome evolution for cancer medicine. *N. Engl. J. Med.* **368**, 842–851 (2013).
46. Buttitta, F. *et al.* Effective assessment of *egfr* mutation status in bronchoalveolar lavage and pleural fluids by next-generation sequencing. *Clin. Cancer Res.* **19**, 691–698 (2013).
47. Kinde, I. *et al.* Evaluation of DNA from the papanicolaou test to detect ovarian and endometrial cancers. *Sci. Transl. Med.* **5**, 167ra164 (2013).
48. Maheswaran, S. *et al.* Detection of Mutations in *EGFR* in Circulating Lung-Cancer Cells. *N. Engl. J. Med.* **359**, 366–377 (2008).
49. Leary, R. J. *et al.* Detection of chromosomal alterations in the circulation of cancer patients with whole-genome sequencing. *Sci. Transl. Med.* **4**, 162ra154 (2012).
50. Chan, K. C. A. *et al.* Cancer genome scanning in plasma: detection of tumor-associated copy number aberrations, single-nucleotide variants, and tumoral heterogeneity by massively parallel sequencing. *Clin. Chem.* **59**, 211–224 (2013).
51. Forshew, T. *et al.* Noninvasive identification and monitoring of cancer mutations by targeted deep sequencing of plasma DNA. *Sci. Transl. Med.* **4**, 136ra168 (2012).
52. Dawson, S.-J. *et al.* Analysis of circulating tumor DNA to monitor metastatic breast cancer. *N. Engl. J. Med.* **368**, 1199–1209 (2013).
- References 51 and 52 used targeted sequencing of circulating tumour DNA to monitor tumour dynamics in patients with advanced solid tumours.**
53. Cristofanilli, M. *et al.* Circulating tumor cells, disease progression, and survival in metastatic breast cancer. *N. Engl. J. Med.* **351**, 781–791 (2004).
54. Murta, M. *et al.* Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. *Nature* **497**, 108–112 (2013).
- This study shows that exome sequencing of circulating tumour DNA can be used to track genomic evolution of metastatic cancers in response to therapy.**
55. Lipson, D. *et al.* Identification of new *ALK* and *RET* gene fusions from colorectal and lung cancer biopsies. *Nature Med.* **18**, 382–384 (2012).
56. Beltran, H. *et al.* Targeted next-generation sequencing of advanced prostate cancer identifies potential therapeutic targets and disease heterogeneity. *Eur. Urol.* **63**, 920–926 (2013).
57. Swanton, C. Intratumor heterogeneity: evolution through space and time. *Cancer Res.* **72**, 4875–4882 (2012).
58. CLIA. Medicare, Medicaid and CLIA programs: regulations implementing the clinical laboratory improvement amendments of 1988. *Fed. Regist.* **57**, 7002–7186 (1992).
59. Wagle, N. *et al.* High-throughput detection of actionable genomic alterations in clinical tumor samples by targeted, massively parallel sequencing. *Cancer Discov.* **2**, 82–93 (2012).
60. Bedard, P. *et al.* Princess Margaret Cancer Centre (PMCC) Integrated Molecular Profiling in Advanced Cancers Trial (IMPACT) using genotyping and targeted next-generation sequencing (NGS). *J. Clin. Oncol.* **31**, 11002 (2013).
61. Miller, V. *et al.* Use of next-generation sequencing (NGS) to identify actionable genomic alterations (GA) in diverse solid tumor types: The Foundation Medicine (FM) experience with 2,200+ clinical samples. *J. Clin. Oncol.* **31**, 11020 (2013).
- References 59–61 show that next-generation sequencing technology can be applied to stored formalin-fixed, paraffin-embedded tumour samples in a clinical setting.**
62. Gargis, A. S. *et al.* Assuring the quality of next-generation sequencing in clinical laboratory practice. *Nature Biotechnol.* **30**, 1033–1036 (2012).
63. Tran, B. *et al.* Feasibility of real time next generation sequencing of cancer genes linked to drug response: results from a clinical trial. *Int. J. Cancer* **132**, 1547–1555 (2013).
64. Dancey, J. E., Bedard, P. L., Onetto, N. & Hudson, T. J. The genetic basis for cancer treatment decisions. *Cell* **148**, 409–420 (2012).
65. Chapman, P. B. *et al.* Improved survival with vemurafenib in melanoma with *BRAF* V600E mutation. *N. Engl. J. Med.* **364**, 2507–2516 (2011).
66. Kopetz, S. *et al.* PLX4032 in metastatic colorectal cancer patients with mutant *BRAF* tumors. *J. Clin. Oncol.* **28**, 3534 (2010).
67. Barretina, J. *et al.* The cancer cell line encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603–607 (2012).
68. Garnett, M. J. *et al.* Systematic identification of genomic markers of drug sensitivity in cancer cells. *Nature* **483**, 570–575 (2012).
69. Papillon-Cavanagh, S. *et al.* Comparison and validation of genomic predictors for anticancer drug sensitivity. *J. Am. Med. Assoc.* **20**, 597–602 (2013).
70. Begley, C. G. & Ellis, L. M. Drug development: raise standards for preclinical cancer research. *Nature* **483**, 531–533 (2012).
71. Sleijfer, S., Bogaerts, J. & Siu, L. L. Designing transformative clinical trials in the cancer genome era. *J. Clin. Oncol.* **31**, 1834–1841 (2013).
72. Dienstmann, R. *et al.* Molecular profiling of patients with colorectal cancer and matched targeted therapy in phase I clinical trials. *Mol. Cancer Ther.* **11**, 2062–2071 (2012).
73. Tsimberidou, A.-M. *et al.* Personalized medicine in a phase I clinical trials program: the MD Anderson Cancer Center Initiative. *Clin. Cancer Res.* **18**, 6373–6383 (2012).
- This study demonstrates that matching investigational treatment to genotype results in better response and survival.**
74. Rodón, J. *et al.* Molecular prescreening to select patient population in early clinical trials. *Nature Rev. Clin. Oncol.* **9**, 359–366 (2012).
75. Hollebecque, A. *et al.* Molecular screening for cancer treatment optimization (MOSCATO 01): a prospective molecular triage trial — interim results. *J. Clin. Oncol.* **31**, 2512 (2013).

76. Kaplan, R. & Brown, L. *MRC FOCUS 4: a New Trial Biomarker-Stratified Phase 2/3 Trial Design* [http://www.methodologyhubs.mrc.ac.uk/annual\\_html\\_network\\_meetings/2013\\_network\\_meeting.aspx](http://www.methodologyhubs.mrc.ac.uk/annual_html_network_meetings/2013_network_meeting.aspx) (Medical Research Council, 2013).
77. Kim, E. S. *et al.* The BATTLE Trial: personalizing therapy for lung cancer. *Cancer Discov.* **1**, 44–53 (2011).
78. Barker, A. D. *et al.* I-SPY 2: an adaptive breast cancer trial design in the setting of neoadjuvant chemotherapy. *Clin. Pharmacol. Ther.* **86**, 97–100 (2009).
79. Berry, D. A., Herbst, R. S. & Rubin, E. H. Reports from the 2010 Clinical and Translational Cancer Research Think Tank Meeting: design strategies for personalized therapy trials. *Clin. Cancer Res.* **18**, 638–644 (2012).
80. Lillie, E. O. *et al.* The n-of-1 clinical trial: the ultimate strategy for individualizing medicine? *Per. Med.* **8**, 161–173 (2011).
81. Gabler, N. B., Duan, N., Vohra, S. & Kravitz, R. L. N-of-1 trials in the medical literature: a systematic review. *Med. Care* **49**, 761–768 (2011).
82. Mick, R., Crowley, J. J. & Carroll, R. J. Phase II clinical trial design for noncytotoxic anticancer agents for which time to disease progression is the primary endpoint. *Control. Clin. Trials* **21**, 343–359 (2000).
83. Kurzrock, R. & Taberner, J. *WINETHER — WIN Therapeutics Clinical Trial*. <http://www.winconsortium.org> (WIN Consortium, 2012).
- This describes an ongoing prospective clinical trial that selects targeted treatment for patients with advanced solid tumours based upon DNA, RNA and microRNA profiling.**
84. Wu, X. *et al.* Clonal selection drives genetic divergence of metastatic medulloblastoma. *Nature* **482**, 529–533 (2012).
85. Tao, Y. *et al.* Rapid growth of a hepatocellular carcinoma and the driving mutations revealed by cell-population genetic analysis of whole-genome data. *Proc. Natl Acad. Sci. USA* **108**, 12042–12047 (2011).
86. Albiges, L., Gerlinger, M. & PREDICT consortium. The PREDICT (Personalised RNA Interference to Enhance the Delivery of Individualised Cytotoxic and Targeted Therapeutics) approach to biomarker discovery in renal cell carcinoma. *Cancer Res.* **72**, S1746 (2012).
87. Gerlinger, M. *et al.* Intratumour heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* **366**, 883–892 (2012).
- This study demonstrates mutational intratumour heterogeneity within spatially separated regions of the same tumour and matched primary and metastatic lesions separated by time.**
88. Tolmachev, V., Stone-Elander, S. & Orlova, A. Radiolabelled receptor-tyrosine-kinase targeting drugs for patient stratification and monitoring of therapy response: prospects and pitfalls. *Lancet Oncol.* **11**, 992–1000 (2010).
89. Gasch, C. *et al.* Heterogeneity of epidermal growth factor receptor status and mutations of *KRAS/PIK3CA* in circulating tumor cells of patients with colorectal cancer. *Clin. Chem.* **59**, 252–260 (2013).
90. Schwarzenbach, H., Hoon, D. S. & Pantel, K. Cell-free nucleic acids as biomarkers in cancer patients. *Nature Rev. Cancer* **11**, 426–437 (2011).
91. Jang, M. H. *et al.* *FGFR1* is amplified during the progression of *in situ* to invasive breast carcinoma. *Breast Cancer Res.* **14**, R115 (2012).
92. Johnson, C. E. *et al.* Identification of copy number alterations associated with the progression of DCIS to invasive ductal carcinoma. *Breast Cancer Res. Treat.* **133**, 889–898 (2012).
93. Heselmeyer-Haddad, K. *et al.* Single-cell genetic analysis of ductal carcinoma *in situ* and invasive breast cancer reveals enormous tumor heterogeneity yet conserved genomic imbalances and gain of *MYC* during progression. *Am. J. Pathol.* **181**, 1807–1822 (2012).
94. The Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70 (2012).
95. Xu, X. *et al.* Single-cell exome sequencing reveals single-nucleotide mutation characteristics of a kidney tumor. *Cell* **148**, 886–895 (2012).
96. Hou, Y. *et al.* Single-cell exome sequencing and monoclonal evolution of a JAK2-negative myeloproliferative neoplasm. *Cell* **148**, 873–885 (2012).
97. Navin, N. *et al.* Tumour evolution inferred by single-cell sequencing. *Nature* **472**, 90–94 (2011).
98. Xu, X. *et al.* Mutational analysis of circulating tumor cells in breast cancer patients by targeted clonal sequencing. *J. Clin. Oncol.* **30**, 10516 (2012).
99. Baselga, J. *et al.* Everolimus in postmenopausal hormone-receptor-positive advanced breast cancer. *N. Engl. J. Med.* **366**, 520–529 (2012).
100. Turner, N. C. & Reis-Filho, J. S. Genetic heterogeneity and cancer drug resistance. *Lancet Oncol.* **13**, e178–e185 (2012).
101. Melcher, C. *et al.* DETECT III — a multicenter, randomized, phase III study to compare standard therapy alone versus standard therapy plus lapatinib in patients with initially HER2-negative metastatic breast cancer but with HER2-positive circulating tumor cells. *Cancer Res.* **72**, SOT1–1–10 (2012).
102. Elisei, R. *et al.* Prognostic significance of somatic RET oncogene mutations in sporadic medullary thyroid cancer: a 10-year follow-up study. *J. Clin. Endocrinol. Metab.* **93**, 682–687 (2008).
103. Park, S. *et al.* Discordance of molecular biomarkers associated with epidermal growth factor receptor pathway between primary tumors and lymph node metastasis in non-small cell lung cancer. *J. Thorac. Oncol.* **4**, 809–815 (2009).
104. Schmid, K. *et al.* *EGFR/KRAS/BRAF* mutations in primary lung adenocarcinomas and corresponding locoregional lymph node metastases. *Clin. Cancer Res.* **15**, 4554–4560 (2009).
105. Kim, H. *et al.* Detection of *ALK* gene rearrangement in non-small cell lung cancer: a comparison of fluorescence *in situ* hybridization and chromogenic *in situ* hybridization with correlation of *ALK* protein expression. *J. Thorac. Oncol.* **6**, 1359–1366 (2011).
106. Janjigian, Y. Y. *et al.* Prognosis of metastatic gastric and gastroesophageal junction cancer by HER2 status: a European and USA International collaborative analysis. *Ann. Oncol.* **23**, 2656–2662 (2012).
107. Bang, Y.-J. *et al.* Trastuzumab in combination with chemotherapy versus chemotherapy alone for treatment of HER2-positive advanced gastric or gastro-oesophageal junction cancer (ToGA): a phase 3, open-label, randomised controlled trial. *Lancet* **376**, 687–697 (2010).
108. Vakiani, E. *et al.* Comparative genomic analysis of primary versus metastatic colorectal carcinomas. *J. Clin. Oncol.* **30**, 2956–2962 (2012).
109. Colombino, M. *et al.* *BRAF/NRAS* mutation frequencies among primary tumors and metastases in patients with melanoma. *J. Clin. Oncol.* **30**, 2522–2529 (2012).
110. Antonescu, C. R. *et al.* Acquired resistance to imatinib in gastrointestinal stromal tumor occurs through secondary gene mutation. *Clin. Cancer Res.* **11**, 4182–4190 (2005).
111. Heinrich, M. C. *et al.* Molecular correlates of imatinib resistance in gastrointestinal stromal tumors. *J. Clin. Oncol.* **24**, 4764–4774 (2006).
112. Walker, I. *et al.* Utilizing a collaborative working model to optimize molecular analysis of solid tumors in the Cancer Research UK's stratified medicine programme. *J. Clin. Oncol.* **31**, 11094 (2013).
113. Rollins, B. *et al.* PROFILE: Broadly based genomic testing for all patients at a major cancer center. *J. Clin. Oncol.* **31**, 1531 (2013).
114. Sequist, L. V. *et al.* Implementing multiplexed genotyping of non-small-cell lung cancers into routine clinical practice. *Ann. Oncol.* **22**, 2616–2624 (2011).
115. MD Anderson Cancer Center. *The Sheik Khalifa Bin Zayed Al Nayan Institute for Personalized Cancer Therapy* <http://www.mdanderson.org/education-and-research/research-at-md-anderson/personalized-advanced-therapy/institute-for-personalized-cancer-therapy/index.html> (MD Anderson Cancer Center, 2012).
116. Mendelson, J. Personalizing oncology: perspectives and prospects. *J. Clin. Oncol.* **31**, 1904–1911 (2013).
117. Voest, E. *Delivering Stratified Medicine: a Strategic Overview*. [http://www.eacr.org/mpathcourse2012/speaker\\_presentations.php](http://www.eacr.org/mpathcourse2012/speaker_presentations.php) (EACR, 2012).
118. Callaway, E. Norway to bring cancer gene tests to the clinic. *Nature* <http://dx.doi.org/10.1038/nature.2012.9949> (2 February, 2012).
119. Vall d'Hebron Institute of Oncology. *Cancer Genomics Group*. [http://www.vhio.net/research/core-technologies/cancer-genomics-group/en\\_index.php](http://www.vhio.net/research/core-technologies/cancer-genomics-group/en_index.php) (Vall d'Hebron Institute of Oncology, 2010).
120. Vanderbilt-Ingram Cancer Center. *Personalized Cancer Therapy at Vanderbilt-Ingram Cancer Center*. <http://www.vicc.org/personalized> (Vanderbilt-Ingram Cancer Center, 2011).
121. Garraway, L. A. *Genomics-Driven Clinical Trials in Oncology: principles and practice*. <http://www.iom.edu/activities/disease/NCPP/2013-Feb-11.aspx> (Institute of Medicine, 2013).
122. Getz, M. & Boughey, J. *Breast Cancer Genome Guided Therapy Study (BEAUTY)*. <http://mayoresearch.mayo.edu/center-for-individualized-medicine/breast-cancer-study.asp> (Mayo Clinic, 2012).
123. Kummer, S. *Challenges in Incorporating Integral NGS into Early Clinical Trials*. [http://cdp.cancer.gov/next-generation\\_agenda.htm](http://cdp.cancer.gov/next-generation_agenda.htm) (National Cancer Institute, 2012).
124. Williams, P. *Request for Information (RFI): Efficient Implementation of a Sequencing Network for the Proposed NCI-MATCH Clinical Trial*. <http://www.fdbdo.com/rfi13-151/> (Fort Detrick Business Development Office, 2013).

**Acknowledgements** Supported in part by the Cancer Care Ontario Applied Cancer Research Units Grant (P.L.B., L.L.S.) and by the US National Institute of Health Grant U01 GM61393 (M.J.R.).

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare competing financial interests: details accompany the full-text HTML version of this paper at [go.nature.com/jofa4v](http://go.nature.com/jofa4v). Readers are welcome to comment on the online version of this article at [go.nature.com/jofa4v](http://go.nature.com/jofa4v). Correspondence should be addressed to L.L.S. ([lillian.siu@uhn.ca](mailto:lillian.siu@uhn.ca)).



# Selection and adaptation during metastatic cancer progression

Christoph A. Klein<sup>1,2</sup>

**Cancer is often regarded as a process of asexual evolution driven by genomic and genetic instability. Mutation, selection and adaptation are by convention thought to occur primarily within, and to a lesser degree outside, the primary tumour. However, disseminated cancer cells that remain after 'curative' surgery exhibit extreme genomic heterogeneity before the manifestation of metastasis. This heterogeneity is later reduced by selected clonal expansion, suggesting that the disseminated cells had yet to acquire key traits of fully malignant cells. Abrogation of the cells' progression outside the primary tumour implies new challenges and opportunities for diagnosis and adjuvant therapies.**

Concern that progress in treating patients with epithelium- or neuroectoderm-derived cancers is simply too slow is increasing. Notwithstanding the great enthusiasm about the potential for improvements to cancer treatment that was evoked by cancer genome projects and the cancer stem-cell concept in the past decade, targeted therapies currently prolong the lives of patients with metastasis by only a few weeks or months<sup>1–4</sup>. Moreover, they often have little or no benefit when given to at-risk patients with no metastasis, that is, as adjuvant therapy<sup>5–7</sup>. In light of these observations and the multilevel complexity and heterogeneity of systemic cancer in particular, focusing merely on individual genetic alterations and corresponding targeted therapies is unlikely to be the most promising approach. We have learned and will probably continue to learn that many, if not all, genes have roles in certain subtypes of some cancers for particular steps in the metastatic cascade in some models. But can we use this knowledge to improve patient treatment?

Since its description by Cairns and Nowell<sup>8,9</sup>, the evolutionary concept of cancer has become widely accepted; this concept has recently been updated in an excellent review<sup>10</sup>, and the molecular mechanisms of metastasis have been reviewed and conceptualized in detail<sup>11–13</sup>. One limitation of the literature on metastasis is that findings are readily included without acknowledgement of the relevance and limitations of patient- and model-derived data. However, mutation, selection and adaptation are linked to the environment of the evolving cancer cell; therefore, differences in these are likely to matter. This Perspective concentrates on human metastasis by attempting to derive an evolutionary concept of cancer based on patient-derived molecular and clinical data, focusing on the early evolution of systemic cancer. However, in doing so, it comes with important caveats, the most relevant being the paucity of information about the phenotypes and early interactions of cancer cells that spread to distant sites. Once these cells form a metastatic colony, the tumour-generated environments of metastases (whether in the lungs, liver or elsewhere) resemble, at least morphologically, those of primary tumours. Consequently, the dynamics of cellular adaptation resulting from heritable genetic or epigenetic changes or phenotypic plasticity (see the Review by Meacham and Morrison on page 328) on are poorly understood in patients. This Perspective summarizes what can be deduced from currently available, mostly genomic, data and may stimulate research into the emergence of cellular cancer phenotypes that cause lethal disease.

## Hallmarks of benign lesions and malignant cancer

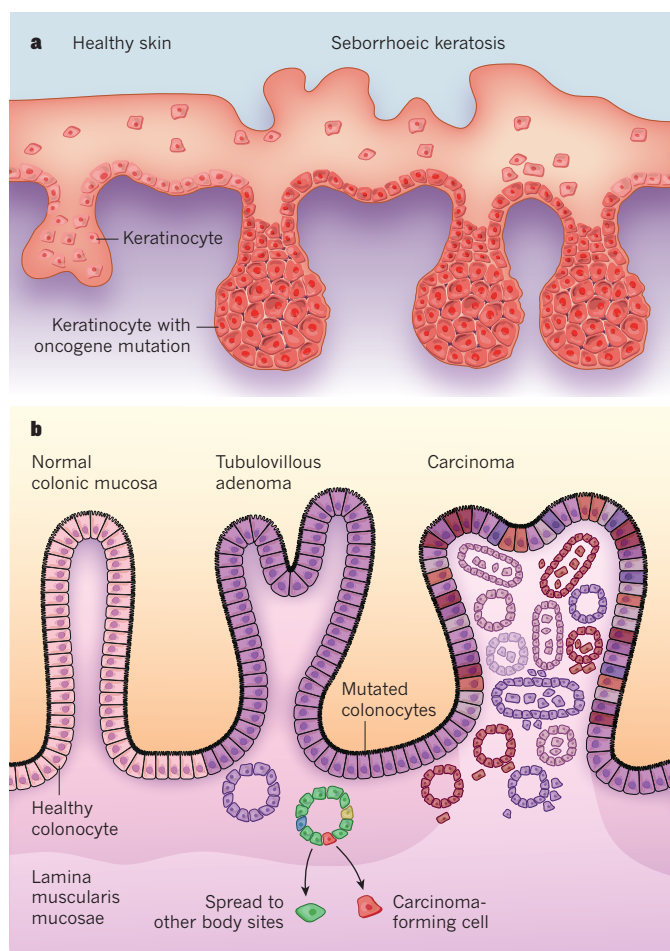
It might be helpful to recall the basic characteristics of cancer before considering systemic progression in detail. Malignant epithelial tumours are able to form vascularized colonies (metastases) at different sites in the body, whereas benign tumours cannot. By definition, then, cancers can invade and metastasize. However, the molecular differences between benign and malignant tumours are woefully understudied. Apart from circumstantial investigations of their genetics, we have a relatively poor molecular appreciation of benign tumours.

Benign tumours in principle can be divided into two groups: those with a statistical association with progressing to cancer and those without. In many instances, the risk of progression is not known, because proving the transition from benign to malignant is not easy. A benign lesion under the pathologist's microscope will never progress to cancer, so the lesion's history and fate can only be linked by association with cancer detected simultaneously or post hoc in the same patient. It is illuminating to consider examples in which malignant conversion is thought not to occur, in which a biological continuum of benign and malignant cancers is debatable and in which it is widely accepted.

The most insightful and comprehensive studies yet on benign lesions that apparently have no risk of malignant conversion have come from Hafner, Real and their teams. From human benign skin tumours, such as seborrheic keratosis one of the most common<sup>14,15</sup>, they identified multiple mutations of oncogenes such as *FGFR3*, *PIK3CA*, *KRAS*, *EGFR*, *HRAS* and *AKT1* (listed in order of decreasing frequency), demonstrated activation of the mutated signalling pathways and showed that no senescence program is activated by these mutations<sup>15</sup>. Of note, *FGFR3* kinase domain mutations leading to the highest constitutive kinase activity<sup>16</sup> are more frequent in benign seborrheic keratosis than in urothelial carcinoma, paradoxically suggesting that just benign lesions may harbour mutations with strong oncogenic potential. Although only a limited number of candidate oncogenes were tested, combinations (most frequently (42%) *FGFR3* and *PIK3CA*) were uncovered, and it is reasonable to expect that genome-wide analyses will reveal additional combinations. There were no signs of genome instability as tested by array comparative genomic hybridization (CGH)<sup>15</sup> (Fig. 1a).

In breast cancer, the debate on the existence of a biological continuum from benign to malignant lesions is ongoing<sup>17,18</sup>. The true relationship of ductal hyperplasia, atypical hyperplasia and ductal carcinoma *in situ* (DCIS) is unclear. This may partially reflect the relative difficulty of

<sup>1</sup>Experimental Medicine and Therapy Research, University of Regensburg, Regensburg, Germany. <sup>2</sup>Personalized Tumor Therapy Project Group, Fraunhofer Institute for Toxicology and Experimental Medicine, Regensburg, Germany.



**Figure 1 | Mutations in benign tumours.** Proving the transition from benign lesion to malignancy is not easy. It can be helpful to consider examples in which conversion is thought not to occur and in which the continuum between benign and malignant cells is accepted. **a**, Seborrheic keratosis is a benign skin tumour in which proliferative cells harbour various mutated oncogenes such as *FGFR3* and *PIK3CA*, but in which neither conversion to malignancy nor genome instability is observed. **b**, Clonal diversity is generated during the adenoma–carcinoma sequence of colonic dysplasias. Genome diversity (shown by the differently coloured cells) seems to emerge at the transition from dysplasia to carcinoma. Cells may migrate beyond the lamina muscularis mucosae and form a carcinoma or spread to distant sites in the body. Note, however, that unlike *in situ* carcinomas of the breast, cancer cell dissemination has not yet been formally shown for colonic dysplasias.

morphological assessment compared with, for example, lesions in the colon. The few data there are available for benign proliferative breast disease suggest that although hyperplasia may display limited allelic losses, DCIS is principally distinguished by the emergence of genome instability, as delineated by fluorescence *in situ* hybridization (FISH) and CGH<sup>18,19</sup>.

The best-studied example of a putative biological benign–malignant continuum is the adenoma–carcinoma sequence in the colon<sup>20</sup>, and clinical procedures have been developed to link the detection of adenomas with prevention or early diagnosis of colorectal cancer<sup>21</sup>. Malignant lesions are characterized by invasion of the lamina muscularis mucosae and beyond. Initial analysis of small adenomas, large adenomas and carcinomas has uncovered an accumulation of genetic alterations paralleling clinical progression<sup>20</sup>. The two main classes of benign lesion are conventional adenomas (comprising tubular, tubulovillous and villous adenomas) and serrated polyps (comprising hyperplastic polyps, traditional serrate adenomas, sessile serrated adenomas and mixed hyperplastic adenomatous polyps)<sup>22,23</sup>. Molecular analyses have focused on three types of event: point mutations, chromosomal rearrangements and

CpG methylation changes<sup>24</sup>. For conventional adenomas, high-grade dysplasia is related to larger adenoma size and a villous component and is an important risk factor for metastasizing cancer. In molecular terms, high-grade dysplasia correlates with the global frequency of allelic imbalances (Fig. 1b), whereas 5'-methylcytosine changes and *KRAS* mutations can also be found in low-grade dysplasias<sup>25</sup>. For serrated polyps<sup>23</sup>, the hyperplastic polyp accounts for 25–30% of resected large intestinal polyps and is the most frequent category. It has an insignificant malignant potential (under current guidelines, no additional surveillance is required<sup>21</sup>) but frequent *BRAF* or *KRAS* mutations<sup>22,23</sup>. Hyperplastic polyps have a normal karyotype with a small subset displaying simple chromosomal aberrations<sup>26</sup>.

In summary, benign lesions share DNA methylation changes, restricted loss of heterogeneity and multiple oncogenic mutations with malignant lesions, but they are insufficient to predispose lesions to invasion and metastasis. Rather, the difference between benign and malignant lesions is apparently associated with the emergence of genome instability. I therefore suggest that the cardinal hallmark of cancer is an ongoing production of genomic diversity, which enables gene networks to self-organize and individual cells to progress towards metastasis in a process of clonal evolution<sup>10,27</sup> (see Review by Swanton and colleagues on page 338). The extent of genome diversity varies but is usually substantial in solid epithelial cancers and melanoma<sup>28,29</sup>. To elaborate this concept, important data are still needed, including direct genome-wide cell-by-cell comparison between genomic diversity in benign lesions and early invasive cancers and careful analysis of global epigenetic change, because epigenetic diversity may also drive clonal evolution and influence genomic stability<sup>10</sup>.

### Cellular diversity and metastasis

The importance of diversity in malignancy is further supported by a study in which Maley and colleagues demonstrated that a high index of clonal diversity predicts progression from a pre-malignant condition (in this case the progression of dysplastic Barrett's oesophagus) to an aggressive cancer (adenocarcinoma of the oesophagus)<sup>30</sup>. Such a direct clinical link between diversity and invasive or metastasizing cancers is lacking for other cancers. However, in another study, whereas clonal diversity between DCIS and invasive breast cancer was similar, breast cancer subtypes displayed different degrees of diversity<sup>31</sup>. The less aggressive (that is, less metastatic<sup>32</sup>) luminal A cancers comprised a few dominant cancer cell populations, whereas the more aggressive basal-like and HER2<sup>+</sup> cancers contained a wider array of less abundant tumour cell types<sup>31</sup> more frequently, when specific genomic loci were investigated by FISH.

If cellular diversity correlates with benign and malignant states, and if increasing diversity indices are associated with the propensity to seed colonies at ectopic sites, what is the role of diversity in metastasis? An obvious hypothesis is that the generation of variant cells increases the probability that some are sufficiently equipped to successfully migrate to, colonize and survive in distant sites. Typical metastatic sites, such as bone marrow, lung, liver and brain, form heterogeneous environments, which are more likely to provide hostile niches for intruding cancer cells than permissive ones. Disseminating cell population size, the proportion of cells able to self-renew, their genotypic diversity and phenotypic plasticity (the ability of one genotype to elicit more than one phenotype in different environments) are all determinants of metastatic success.

What follows is a discussion of disseminated cancer cell (DCC) adaptation to ectopic sites (those outside the site of origin). This, it is argued, is an evolutionary process engendering fitness to generate a metastatic colony, requiring researchers to consider whether selection occurs within or outside the primary tumour.

### DCC genomes in early and advanced systemic cancer

Cancer cells often start to disseminate on generation of genetic diversity at the primary site. Early dissemination can be deduced from disease course and patient-derived data<sup>33–35</sup> (reviewed in ref. 35). Direct

evidence is growing that presumptive pre-invasive lesions such as *in situ* carcinomas can seed cancer cells<sup>36–38</sup> (Fig. 1b).

One might imagine that dissemination forms a bottleneck that restrains genome diversity found among individual DCCs. If this is so, we have not yet identified the underlying rules, suggesting that these are either too subtle or complex, or that dissemination is linked to a specific phenotype, not genotype. Note, however, that we are lacking sufficient data about the phenotype of DCCs. Genomic cell-by-cell analysis of DCCs isolated from bone marrow or lymph nodes of patients with carcinoma but without overt metastases revealed substantial heterogeneity during occult (undetectable by clinical imaging) systemic spread before manifestation of clinical metastasis<sup>39,43</sup> (Fig. 2a, c). When more than one cell was analysed, sibling cells from an individual patient rarely shared multiple chromosomal aberrations, but frequently possessed distinctive genomes<sup>39,41,43</sup>. Thus, genome diversity is evident in early systemic cancer. DCCs from bone marrow and lymph nodes have also been shown to harbour characteristically different chromosomal aberrations<sup>42</sup>.

In all cancers investigated, DCC genomes become highly similar once metastases have clinically manifested, suggesting that expansion of an aggressive clone parallels the development of generalized, incurable disease<sup>39,43</sup> (Fig. 2a, b). The pattern of chromosomal aberrations in late-stage disease seems relatively stable, because, even after several cycles of high-dose chemotherapy, a shared common baseline of gains and losses can be identified<sup>39</sup> (Fig. 2c).

### Macroevolution and microevolution of DCC genomes

DCCs in patients with metastases present a recurrent pattern of genomic and genetic changes referred to here as an attractor<sup>44</sup>. This preferred genomic state is also seen in circulating tumour cells (CTCs) isolated from patients with metastasis. The few CTC data so far mostly describe genome alterations of pooled CTCs<sup>45,46</sup>, but detection of typical late-stage alterations is evident. Putative attractor genomes of DCCs and CTCs from patients with manifest metastasis resemble those found in corresponding primary tumours<sup>39,41,43,47</sup>.

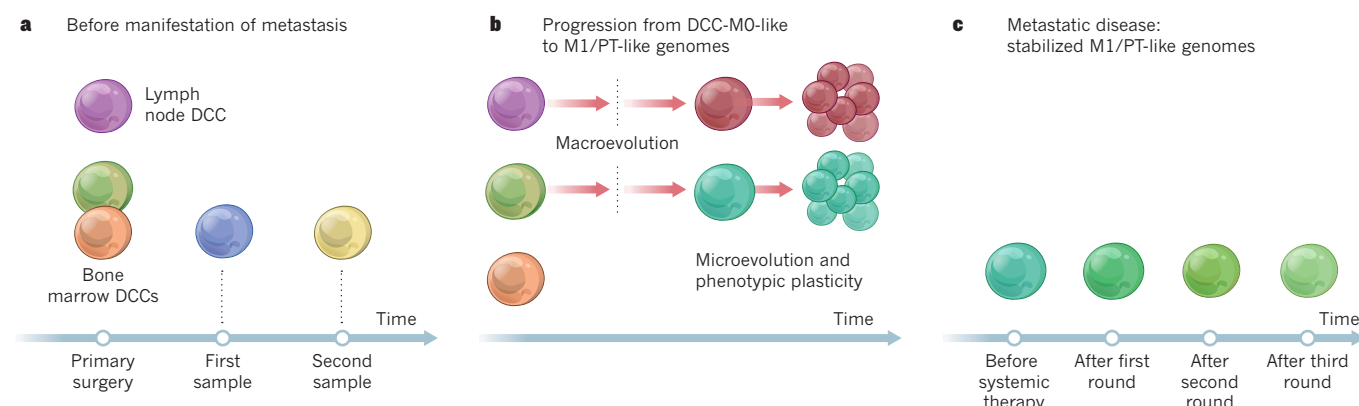
The genomic difference between DCCs isolated from patients with no distant cancer spread (M0 stage) and patients with cancer spread (M1 stage) is striking and fundamental (Fig. 2). Without knowing the disease stage of a patient with breast cancer, it is possible to predict with 85% accuracy whether the patient has manifest metastasis<sup>41</sup> by determining the chromosomal aberrations of single cytokeratin-positive cells isolated from the bone marrow. This is not possible with genomic information from the primary tumour. For the remainder of the Perspective, I will therefore differentiate between DCC-M0-like and M1/primary tumour

(PT)-like genomes.

DCC-M0-like genome data are limited and await comprehensive investigation with modern sequencing technologies. So far, the number of chromosomal changes has been found to be lower than in M1/PT-like genomes<sup>39,41,43</sup>; typical chromosomal aberrations expected for the cancer under consideration are mostly absent<sup>41,43</sup>; evidence for telomere crisis is lacking as usually gains or losses of whole chromosomes, but less frequently chromosome breaks, occur<sup>41</sup>; point mutations characteristic of the cancer under consideration are mostly absent<sup>39</sup>; and subchromosomal losses and gains often precede the accumulation of gains or losses of whole chromosomes<sup>40</sup> (Fig. 2a). Thus, DCCs isolated mostly from the bone marrow of M0-stage patients seem to disseminate in a genomically ‘immature’ state in which they have not yet acquired typical mutations in oncogenes and tumour suppressor genes or copy number changes, and in some cases may not be immortal — as defined by the absence of signs of passage through telomere crisis. The genomes of such cells will be referred to as DCC-M0-like, although individual DCCs in a patient may exhibit these five traits of malignancy to different degrees.

In contrast to DCC-M0-like genomes, M1/PT-like genome changes match the chromosomal signatures of their associated cancers. Evidence that typical changes resemble attractor states in the somatic evolution of a cancer comes from many studies such as cytogenetic and molecular genetic analyses that reveal cancers acquire chromosomal aberrations non-randomly, with preferred patterns of gain and loss for each cancer<sup>29,48</sup>. For example, in breast cancer, typical M1/PT-like aberrations include losses of chromosome 8p and 17p and gains of 8q and 17q, which are mostly absent in DCC-M0-like genomes. But clearly, similarities between cancers of one type from different patients do not imply the cancers are clonal and, accordingly, the attainment of similar attractors does not imply clonal descent, for example, of metastasis from a specific region of a primary tumour. Furthermore, sequencing studies revealing genomic heterogeneity in primary tumours and metastases suggest that several attractor states may exist for a given cancer<sup>49</sup>.

It may be helpful to differentiate between macroevolution and microevolution in the somatic evolution of DCCs. Here, macroevolution refers to the transition of relatively normal DCC-M0-like genomes to M1/PT-like genomes (Fig. 2b), and microevolution to the ongoing evolution of DCCs with M1/PT-like genomes. Macroevolution results from evolutionary shifts<sup>50</sup>, and it is difficult with our current understanding to explain the transition from DCC-M0-like to M1/PT-like genomes without such a shift. Evolutionary shifts could be induced by processes such as telomere crisis or the inactivation of an important tumour suppressor (for example, p53). Atypical chromosomal changes found in the diverse



**Figure 2 | Macroevolution and microevolution of breast cancer genomes.** **a**, Heterogeneity of genomes of disseminated cancer cells (DCCs) before manifestation of metastasis (DCC-M0-like). Bone marrow and lymph node DCCs have different genomes, as do DCCs isolated at different time points. DCC-M0-like genomes have few and atypical copy number alterations, no signs of telomere crisis, subchromosomal loss that occurs before whole chromosome change and few point mutations. **b**, Progression of DCC-M0-like genomes to the genomes of cells that are found in manifest

metastases/primary tumours (M1/PT-like). Cells with DCC-M0-like genomes must undergo macroevolution to M1/PT-like genomes. Additional layers of diversity are generated by clonal expansion, microevolution and phenotypic plasticity. **c**, Relative stability of M1/PT-like genomes after manifestation of metastasis. Even over the course of several rounds of systemic treatment genomes display similarity. M1/PT-like genomes have many and typical copy number alterations, chromosomal breaks, typical point mutations and evidence of clonal expansion.



DCC-M0-like genomes may increase the fitness of individual cells in a specific micro-environment. Chromosomal changes alter the expression of many genes<sup>51</sup>, potentiating progression to M1/PT-like genomes. Once M1/PT-like attractor states have been attained, the expanding metastasis-forming cell populations may evolve gradually. Microevolution is this gradual evolution within a population of cells that have reached a genomic attractor state.

### Plausibility and potential relevance of macroevolution

DCC macroevolution is unstudied but may be important for several reasons. First, unlike benign tumours or putative precursor lesions, DCC-M0-like genomes belong to a DCC population directly linked to cancer through an originating primary tumour. In its defining ability to spread is the potential seed of metastasis. Because DCCs represent the earliest stages of cancer, they may allow more clinically relevant investigation of carcinogenesis than benign or putative precursor lesions, for which transition to malignancy is unpredictable. Second, high-resolution DCC-M0-like genome analysis may help to identify early changes and therapy targets<sup>40</sup>. Third, the transition from DCC-M0-like genomes to the attractor state of M1/PT-like genomes may form a significant hurdle, which could be raised by new modes of therapeutic intervention, or may suggest new barriers that could be induced artificially.

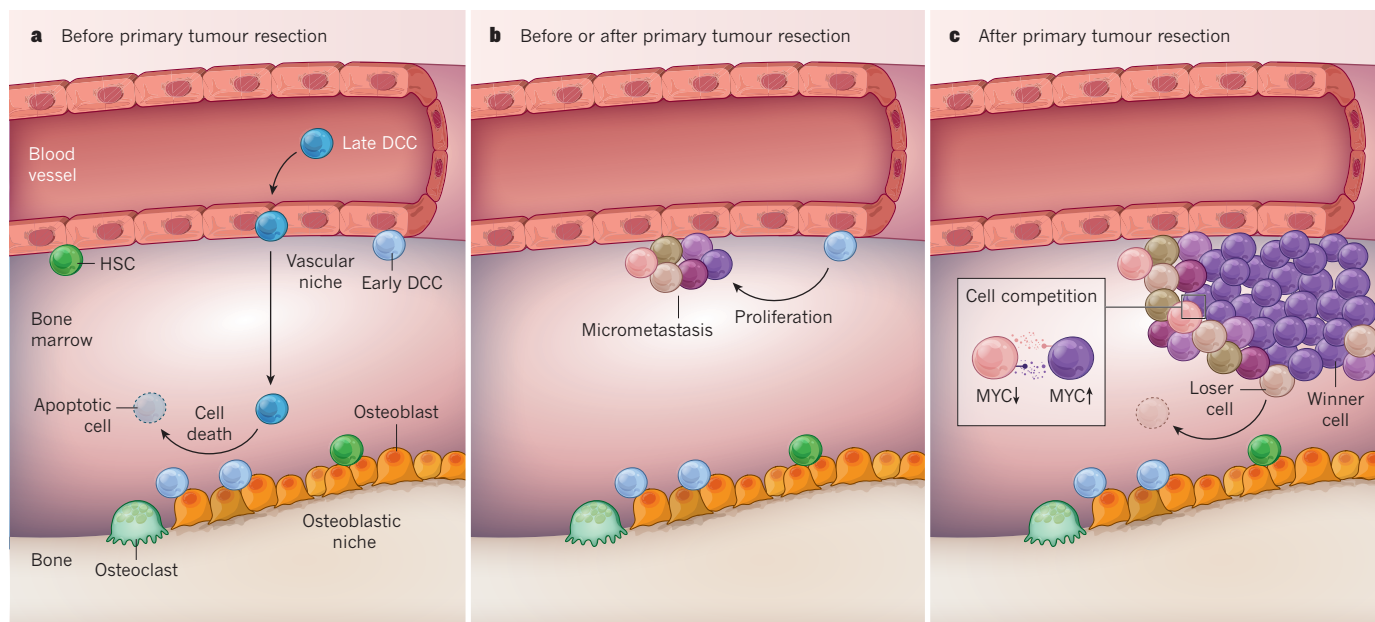
Although the occurrence of early dissemination and the existence of DCCs with immature DCC-M0-like genomes are increasingly accepted, the relevance of these cells as seeds of metastasis is questioned and negated<sup>12,13</sup>. It may be impossible to prove the cells' relevance in patients because the founder cell of a metastasis may never be identified and DCC-M0-like genomes may be outcompeted and disappear once M1/PT-like genomes emerge. Support for the cells' role in progression of human cancer will come from a consistent and coherent clinical picture if it exists. However, circumstantial evidence supporting the argument for DCC-M0-like genomes already exists.

### Rare detection of M1/PT-like DCCs in M0-stage patients

A major argument in favour of a role for DCCs with DCC-M0-like genomes is that it is exceptional for DCCs with M1/PT-like genomes to be detected in bone marrow at the time of curative surgery. DCCs

comprise all cancer cells that spread before the removal of the primary tumour. If cells with M1/PT-like genomes form the predominant cell population in the primary tumour at the time of sampling, why are they rare? Technically, detection of DCCs depends on the sensitivity and specificity of available markers. The two most commonly used markers for epithelium-derived DCCs, glandular cytokeratins (mostly keratin 8, 18 and 19) and EpCAM (originally called 17-1A antigen) were introduced by Riethmüller and Schlimok in the 1980s (ref. 52). These markers are suited to the detection of epithelial DCCs in tissues such as blood, bone marrow and lymph nodes because these tissues lack epithelial cells. However, expression of both markers can be low or absent in epithelial cells, so DCCs may escape detection — the reasons for this downregulation or loss are not understood. Both markers are expressed in early embryonic development<sup>53</sup> or human embryonic stem cells<sup>54,55</sup> in normal and malignant cells<sup>56,57</sup> with differentiated, progenitor-cell or stem-cell features<sup>58–61</sup>. Detection of cells with either marker in the blood, bone marrow or lymph node tissue of cancer patients is associated with poor prognosis<sup>62</sup>; however, there are typically only 1–10 positive cells for every 10<sup>6</sup> normal bone marrow or lymph node cells in less than 50% of M0-stage patients. In summary, despite limitations that are inherent to all markers, there is good reason to think cytokeratin and EpCAM expression capture relevant DCCs. Moreover, because both DCCs from metastatic and non-metastatic patients have been identified with the same markers<sup>39–43</sup>, it would be inconsistent and unfounded to suggest that the markers regularly miss cells with M1/PT-like genomes in M0-stage patients but detect them in M1-stage patients.

Biologically, the explanation for M1/PT-like genome rarity is based on studies involving thousands of patients<sup>37,62,63</sup>. Increasing DCC counts did not correlate with tumour size, even though there may have been several hundred-fold more cancer cells in large primary tumours than in small ones, suggesting that constantly disseminating cancer cells in bone marrow do not simply accumulate over years of primary tumour growth<sup>37,62,63</sup>. There are three possible explanations for this. First, there is no increased number of DCCs in bone marrow over time because dissemination and death of DCCs in bone marrow are balanced. Second, early cancers are more capable than late cancers of generating cells that are adapted to disseminate, integrate and survive



**Figure 3 | Early steps of metastasis at ectopic sites in bone marrow.** **a**, Before removal of the primary tumour, cell competition for occupancy of vascular and osteoblastic niches may favour early disseminated cancer cells (DCCs). Both niches may be used by DCCs; however, late DCCs may find that the niches are occupied by early DCCs and die as a result of apoptosis. **b**, Before or after surgery, DCCs must proliferate to form a

micrometastasis, which is essential for selective adaptation and progression towards cells with M1/PT-like genomes. **c**, During micrometastasis formation, cells may engage in competition whereby secreted factors allow the recognition of fitness differences between cells. Cells with high levels of MYC, which have higher fitness (winner cells), induce apoptosis in cells with low MYC, which have lower fitness (loser cells).

in bone marrow. Third, there are a limited number of niches in bone marrow (and probably in other organs), so these are unavailable to cells that disseminate later.

Balanced dissemination and death, can be discarded at least for prostate cancer, as the percentage of DCC-positive bone marrow samples is constant (around 20%) before surgery and more than a decade post-surgery in non-progressing patients<sup>43</sup>. All patients had undergone radical prostatectomy, so no later DCCs could be seeded, suggesting that DCCs are long-lived but — contrary to the study's finding — would have continued to accumulate where the primary tumour remained with the ability to seed. Decreasing dissemination or integration and the limited number of niches, probably both contribute. Cells representing the predominant clone of the primary tumour at resection may rarely disseminate or compete with early DCCs or other cells that occupy the niche and seem to be maladapted to displace early-DCCs from their niche (Fig. 3a). Experimentally, prostate cancer cells compete with haematopoietic stem cells (HSCs) for the endosteal niche in bone marrow, and direct competition assays have revealed that the number of niches is limited and that malignant and normal cells have different (lower and higher, respectively) affinities for them<sup>64</sup>. If these findings can be extended to patients, they may further explain why bone-seeking cancers (such as breast and prostate) are not initially associated with much higher numbers of DCCs than cancers that exhibit bone metastasis less frequently (such as colon cancer)<sup>62</sup>, as early seeding rates would be sufficient to fill the available niches.

### Asynchronous and independent evolution of metastases

For all carcinomas, patients initially present with local lymph node metastases in the absence of distant metastasis far more frequently than vice versa. This is reflected in the classic International Union Against Cancer (UICC) staging system, for which stage III represents local lymph node metastasis without distant metastasis and stage IV represents disease with distant metastasis, regardless of lymph node status. It would therefore be expected, that more DCCs from lymph nodes progress further at the time of curative surgery and contain M1/PT-like genomes more frequently, particularly when micrometastases are present, than do DCCs from bone marrow. This is indeed the case<sup>42</sup>, suggesting that metastatic colony growth is generally associated with evolution towards M1/PT-like genomes — this is, however, asynchronous. Despite their evolutionary head start, lymph node micrometastases are unlikely sources of distant metastases, as evidenced by many clinical studies demonstrating that removal of the micrometastases has no effect on distant metastases<sup>65–69</sup>, supporting the idea that DCCs at different sites evolve into M1/PT-like cells independently.

### Genomic comparisons of primary tumours and metastases

Cells with DCC-M0-like genomes progressing towards M1/PT-like genomes become similar to the primary tumour — but not identical. Genetic disparity and clonal divergence within the primary tumour, and between primary tumours and paired metastases were observed even before the advent of next-generation sequencing technologies (studies up to 2009 have been summarized in ref. 70), but have generally been corroborated since<sup>49,71–74</sup>. Disparity is seen for all types of genetic alteration, including point mutations<sup>49,72–74</sup>, rearrangement breakpoints<sup>75</sup>, areas of chromothripsis<sup>71</sup> (an extensive genome scrambling<sup>76</sup>), collectively supporting ongoing genomic evolution at metastatic sites. However, other studies have found higher similarity between primary tumours and paired metastases or relapsing cancers<sup>77–79</sup>. Explanations for why some primary tumours and metastases are more closely related than others are manifold. Clinical details, such as the time interval between primary tumour resection and biopsy or surgery of the metastasis, or the type and extent of systemic treatment are often omitted. Cancers with explosive growth at primary and distant sites might result from aggressive, highly similar clones<sup>78</sup>, whereas metastases arising after long latency periods may differ extensively<sup>73</sup>. Toxic chemotherapy drugs may homogenize cancer cell populations at primary and distant sites in

stage IV patients who have not undergone surgery, by generating cells with high phenotypic plasticity that are able to colonize the whole body (Fig. 4). This highlights the need for clinical information, without which studies<sup>79</sup> are difficult to evaluate.

### Time considerations

The latency period that is needed to initiate systemic cancer from DCCs with DCC-M0-like genomes would provide a much better explanation of the time courses of patients' diseases than the assumption that the most aggressive fully malignant clones initiate the metastases. Most human cancers double their tumour volume within 60–250 days, but prostate cancer has a tumour volume doubling time (TVDT) of 700 days<sup>35</sup>. Consequently, human cancers can often take several years to diagnose. The fastest growing 5% of breast cancers can reach 1–2 cm within 1 year, whereas the slowest growing 5% reach this size within about 5 decades<sup>80,81</sup>. The time course of metastasis suggests that metastases have similar growth rates to primary tumours<sup>35,80</sup>. By contrast, the TVDT of xenograft models using cell lines, such as MDA-MB-231, is less than 1 day and that of transgenic or knockout mouse models between 3 and 12 days. Since the former start with M1/PT-like genomes and the latter are initiated by transgenes or loss of tumour suppressor genes that result in strong oncogenic potential (changes that are rarely detected in DCC-M0-like genomes), a major proportion of the natural evolution from a normal cell is abrogated, which may at least partially explain DCCs' accelerated growth in models. In short, xenograft and many transgenic cancer models fail to model one or more key phases of human cancer.

### DCC microevolution in advanced metastatic disease

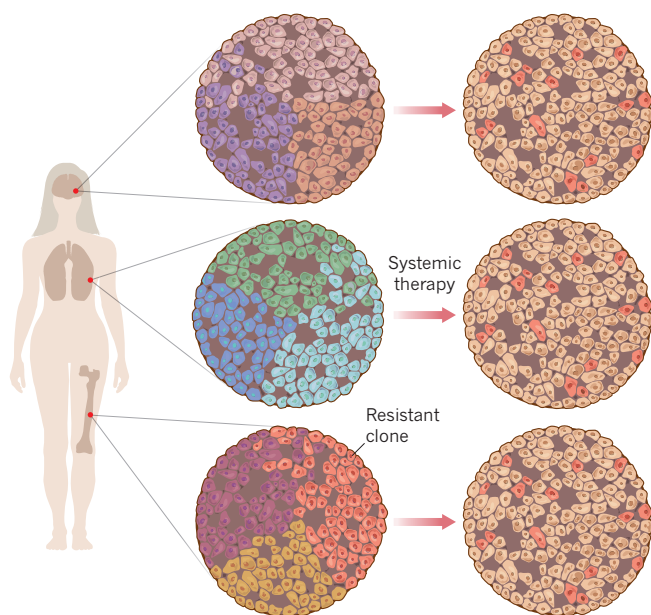
Once metastasis is diagnosed, DCCs and CTCs display M1/PT-like genomes, which display less heterogeneity by CGH analysis. In metastatic disease, microevolution operates as in MDA-MB-231, which have been found to display rather stable M1/PT-like genomes despite selection at different sites, including bone marrow, lung and brain<sup>82–84</sup>. Chromosome profiles for MDA-MB-231 subpopulations that metastasized to different organs were similar despite marked differences in expression patterns of metastasis-associated genes and metastatic activities and the exhibition of site-specific gene expression signatures<sup>83</sup>. Cells with M1/PT-like genomes, once they have undergone several rounds of selection, seem to develop high phenotypic plasticity, enabling them to resist various types of selection pressure. Therefore, metastasis models based on cell lines model metastatic processes that occur in end-stage metastatic patients.

### Factors promoting metastatic progression

If the progression of DCCs with DCC-M0-like genomes is important for human metastasis and if the progression provides a time window for intervention, what are the underlying mechanisms? Answering this may enable new paths for intervention. Currently, there are no data but by analogy to carcinogenesis and *in vitro* selection experiments cell competition and micro-environmental factors may have major roles.

The recognition of DCCs with 'incompletely mutated' DCC-M0-like genomes is reminiscent of studies (for example, see ref. 85) that have attempted to induce cancer from normal cells. These experiments uncovered the need for initiating (genotoxic and mutating) as well as promoting (non-genotoxic and non-carcinogenic) agents<sup>85</sup> to generate skin cancer. Promoting agents, such as 12-O-tetradecanoylphorbol-13-acetate (TPA), can exert their effects long after carcinogens have initiated DNA lesions and been removed<sup>86</sup>. *In vitro* selection experiments have shown that cell density and cell–cell contact are crucial for full spontaneous transformation of initiated (mutated) cells<sup>87–89</sup>. Therefore, single DCCs must first start proliferation (possibly out of dormancy) to form a micrometastasis (Fig. 3b). Then, the driving force for progression of DCC-M0-like genomes towards M1/PT-like genomes may likewise be selection and cell competition (Fig. 3c). Cell competition is a selection process in which 'loser' cells with low fitness are eliminated (for





**Figure 4 | Heterogeneity of metastasis and homogenization by systemic therapies.** Independent somatic evolution, cell competition and various systemic and site-specific selection pressures (such as the extracellular matrix, hypoxia and immunosurveillance) generate metastatic diversity. Although systemic therapy may eliminate cells, resistant clones may overgrow after selection by therapy. These resistant cells may have been present as a minority at one metastatic site before therapy and then seeded other sites in the body (shown here) or they may have already been present at various metastatic sites. The extent of an homogenizing effect by repeated rounds of systemic therapies on resistant cancer cells is currently unknown. Cells that survive multiple rounds may display high phenotypic plasticity for growth throughout the body.

example, through induction of apoptosis) by 'winner' cells with higher fitness<sup>90,91</sup>. During development, we know of two types of competition: adhesion-based competition for niche occupancy and direct cell–cell comparison of metabolic status<sup>90</sup>. In both cases, cell fitness is sensed and communicated, and variation of both could play a part in metastasis. The relevance of cell competition to cancer is increasingly recognized thanks to the work of Moreno<sup>92</sup> and colleagues, and it is noteworthy that competition involves pathways known to have a role in cancer, including MYC, TP53, WNT and JAK–STAT pathways<sup>91</sup>. Selection would produce winning cells with alterations in these pathways and indeed amplification of MYC on chromosome 8q is clearly associated with M1/PT-like genomes in breast and prostate cancer DCCs<sup>41,43</sup>, as is loss of 17p and TP53 mutations<sup>39,41</sup>. Whether competition for anatomical niches has a role in systemic cancer is still debated, but data from animal models suggest that it does<sup>64</sup>. If cell competition is needed to select for higher cell fitness, metastatic progression may occur only if DCCs expand and engage in competition (Fig. 3b, c), because M0-like DCCs left the primary tumour early and thereby avoided cell competition at the primary site.

The tumour-promoting effect of TPA may be due at least partly to altering the micro-environment by inducing massive inflammation<sup>93</sup>. The role of micro-environmental changes (see the Review by Junttila and de Sauvage on page 346) in promoting the transition from DCC-M0-like to M1/PT-like genomes may therefore be of particular interest. Experimental primary tumours may suppress outgrowth of micrometastases by preventing the induction of angiogenesis<sup>94</sup> or by imposing DCC dormancy through secretion of *meta*- and *ortho*-tyrosine<sup>95</sup>. They may also stimulate metastatic outgrowth by the secretion of factors, such as osteopontin or S100A8 that recruit bone-marrow-derived cells to the incipient or established site of metastasis<sup>96–98</sup>. Primary tumour exosome secretions may even evoke the 'education' of tissues such as bone marrow resulting in the long-lasting promotion of metastasis<sup>99</sup>. The local

environment at the distant site, such as sprouting neovasculature, may also trigger outgrowth of dormant DCCs<sup>100</sup>. Although many of these studies have been performed using cell lines with M1/PT-like genomes, there is no reason to think that indirect metastasis-promoting effects through engagement of supportive cell populations would not also affect DCCs with DCC-M0-like genomes.

### Future directions

The concept presented in this Perspective remains to be tested for a broad range of cancer types. Most data are available for breast and prostate cancer and interesting differences between these cancers and more aggressive cancers have been highlighted<sup>101</sup>. The concept should also be scrutinized for its relevance to diagnosis and therapy. It predicts that certain characteristics of M1/PT-like genomes present in DCCs before manifestation of metastasis are predictive of patients with shortened relapse-free survival. It further stresses the need to determine directly the molecular characteristics of DCCs and CTCs so that the evolution of the disease can be monitored at all stages and the appropriate (targeted) therapies selected. Basic research will be needed to address the phenotype of cancer cells adapting to different selective conditions before and after the cancer cells have induced their own tumour-like environment. For therapy development, retarding cellular macroevolution may be much more effective than attempts to kill M1/PT-like cells, and require a change to current adjuvant therapy strategies. Notably, initial attempts have been made specifically to target aneuploid genomes with drugs<sup>102</sup> that are arguably more effective in DCCs with DCC-M0-like genomes that have not yet reached the M1/PT-like attractor state because such drugs target existing aneuploidy-associated stresses<sup>51</sup>. Adjuvant therapy clearly requires substantially different strategies to abrogate metastatic disease<sup>7</sup>, including the design of clinical trials<sup>103</sup> (see Review by Siu and colleagues on page 355). Although evolution of different clones at different sites is unlikely to be synchronized, this should not stop us from exploring options to suppress tumour-promoting effects. Analysis of large clinical and epidemiological data sets for the temporal and spatial distribution of metastases suggests that initially (without interference from genotoxic therapies) neither distant nor lymphatic metastases are themselves able to metastasize<sup>65,104</sup>. Therefore, if we successfully manage to retard progression of DCCs with M0-like genomes to M1/PT-like genomes by non-genotoxic means, more patients may present with surgically manageable metastasis (if they present with any at all), which could provide clinicians with a second chance to completely ablate the patient's disease. ■

Received 19 April; accepted 12 August 2013.

1. Douillard, J. Y. *et al.* Randomized, phase III trial of panitumumab with infusional fluorouracil, leucovorin, and oxaliplatin (FOLFOX4) versus FOLFOX4 alone as first-line treatment in patients with previously untreated metastatic colorectal cancer: the PRIME study. *J. Clin. Oncol.* **28**, 4697–4705 (2010).
2. Mok, T. S. *et al.* Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N. Engl. J. Med.* **361**, 947–957 (2009).
3. Slamon, D. J. *et al.* Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *N. Engl. J. Med.* **344**, 783–792 (2001).
4. Sosman, J. A. *et al.* Survival in BRAF V600-mutant advanced melanoma treated with vemurafenib. *N. Engl. J. Med.* **366**, 707–714 (2012).
5. Alberts, S. R. *et al.* Effect of oxaliplatin, fluorouracil, and leucovorin with or without cetuximab on survival among patients with resected stage III colon cancer: a randomized trial. *J. Am. Med. Assoc.* **307**, 1383–1393 (2012).
6. Goss, G. D. *et al.* A phase III randomized, double-blind, placebo-controlled trial of epidermal growth factor receptor inhibitor gefitinib in completely resected stage IB–IIIA non-small cell lung cancer (NSCLC): NCIC CTG BR.19. *J. Clin. Oncol.* **28** (suppl.), abstr. LBA7005 (2010).
7. Polzer, B. & Klein, C. A. Metastasis awakening: the challenges of targeting minimal residual cancer. *Nature Med.* **19**, 274–275 (2013).
8. Cairns, J. Mutation selection and the natural history of cancer. *Nature* **255**, 197–200 (1975).
9. Nowell, P. C. The clonal evolution of tumor cell populations. *Science* **194**, 23–28 (1976).
10. Greaves, M. & Maley, C. C. Clonal evolution in cancer. *Nature* **481**, 306–313 (2012).

This review summarizes current evolutionary concepts of cancer.



11. Nguyen, D. X., Bos, P. D. & Massague, J. Metastasis: from dissemination to organ-specific colonization. *Nature Rev. Cancer* **9**, 274–284 (2009).
12. Valastyan, S. & Weinberg, R. A. Tumor metastasis: molecular insights and evolving paradigms. *Cell* **147**, 275–292 (2011).
13. Kang, Y. & Pantel, K. Tumor cell dissemination: emerging biological insights from animal models and cancer patients. *Cancer Cell* **23**, 573–581 (2013).
14. Hafner, C. *et al.* Oncogenic *PIK3CA* mutations occur in epidermal nevi and seborrheic keratoses with a characteristic mutation pattern. *Proc. Natl Acad. Sci. USA* **104**, 13450–13454 (2007).
15. Hafner, C. *et al.* Multiple oncogenic mutations and clonal relationship in spatially distinct benign human epidermal tumors. *Proc. Natl Acad. Sci. USA* **107**, 20780–20785 (2010).
- References 14 and 15 provide insight into the genetics and biology of benign skin tumours.**
16. Naski, M. C., Wang, Q., Xu, J. & Ornitz, D. M. Graded activation of fibroblast growth factor receptor 3 by mutations causing achondroplasia and thanatophoric dysplasia. *Nature Genet.* **13**, 233–237 (1996).
17. Allred, D. C. *et al.* Ductal carcinoma *in situ* and the emergence of diversity during breast cancer evolution. *Clin. Cancer Res.* **14**, 370–378 (2008).
18. Boecker, W. *et al.* Ductal epithelial proliferations of the breast: a biological continuum? Comparative genomic hybridization and high-molecular-weight cytokeratin expression patterns. *J. Pathol.* **195**, 415–421 (2001).
19. Chin, K. *et al.* *In situ* analyses of genome instability in breast cancer. *Nature Genet.* **36**, 984–988 (2004).
- This paper suggests that telomere crisis could result in an evolutionary shift at the transition from benign to malignant tumours.**
20. Vogelstein, B. *et al.* Genetic alterations during colorectal-tumor development. *N. Engl. J. Med.* **319**, 525–532 (1988).
21. Winawer, S. J. *et al.* Guidelines for colonoscopy surveillance after polypectomy: a consensus update by the US Multi-Society Task Force on Colorectal Cancer and the American Cancer Society. *Gastroenterology* **130**, 1872–1885 (2006).
22. Carr, N. J., Mahajan, H., Tan, K. L., Hawkins, N. J. & Ward, R. L. Serrated and non-serrated polyps of the colorectum: their prevalence in an unselected case series and correlation of *BRAF* mutation analysis with the diagnosis of sessile serrated adenoma. *J. Clin. Pathol.* **62**, 516–518 (2009).
23. Bettington, M. *et al.* The serrated pathway to colorectal carcinoma: current concepts and challenges. *Histopathology* **62**, 367–386 (2013).
24. Toyota, M. *et al.* CpG island methylator phenotype in colorectal cancer. *Proc. Natl Acad. Sci. USA* **96**, 8681–8686 (1999).
25. Neuville, A. *et al.* Histologic characteristics of non-microsatellite-unstable colon adenomas correlate with distinct molecular patterns. *Hum. Pathol.* **42**, 244–253 (2011).
26. Bardi, G. *et al.* Cytogenetic comparisons of synchronous carcinomas and polyps in patients with colorectal cancer. *Br. J. Cancer* **76**, 765–769 (1997).
27. Heng, H. H. *et al.* Evolutionary mechanisms and diversity in cancer. *Adv. Cancer Res.* **112**, 217–253 (2011).
28. Mitelman, F., Johansson, B. & Mertens, F. *Mitelman Database of Chromosome Aberrations and Gene Fusions in Cancer* <http://cgap.nci.nih.gov/Chromosomes/Mitelman> (NCI, 2013).
29. Teixeira, M. R. & Heim, S. Cytogenetic analysis of tumor clonality. *Adv. Cancer Res.* **112**, 127–149 (2011).
30. Maley, C. C. *et al.* Genetic clonal diversity predicts progression to esophageal adenocarcinoma. *Nature Genet.* **38**, 468–473 (2006).
- This paper established that measures of genetic diversity are a novel form of biomarker.**
31. Park, S. Y., Gonen, M., Kim, H. J., Michor, F. & Polyak, K. Cellular and genetic diversity in the progression of *in situ* human breast carcinomas to an invasive phenotype. *J. Clin. Invest.* **120**, 636–644 (2010).
32. Kennecke, H. *et al.* Metastatic behavior of breast cancer subtypes. *J. Clin. Oncol.* **28**, 3271–3277 (2010).
33. Engel, J. *et al.* The process of metastasis for breast cancer. *Eur. J. Cancer* **39**, 1794–1806 (2003).
34. Klein, C. A. & Holzel, D. Systemic cancer progression and tumor dormancy: mathematical models meet single cell genomics. *Cell Cycle* **5**, 1788–1798 (2006).
35. Klein, C. A. Parallel progression of primary tumours and metastases. *Nature Rev. Cancer* **9**, 302–312 (2009).
36. Banys, M. *et al.* Hematogenous and lymphatic tumor cell dissemination may be detected in patients diagnosed with ductal carcinoma *in situ* of the breast. *Breast Cancer Res. Treat.* **131**, 801–808 (2012).
37. Hüsemann, Y. *et al.* Systemic spread is an early step in breast cancer. *Cancer Cell* **13**, 58–68 (2008).
- This paper established that metastatic dissemination might occur shortly after malignant transformation in transgenic mouse models and patients.**
38. Sängler, N. *et al.* Disseminated tumor cells in the bone marrow of patients with ductal carcinoma *in situ*. *Int. J. Cancer* **129**, 2522–2526 (2011).
39. Klein, C. A. *et al.* Genetic heterogeneity of single disseminated tumour cells in minimal residual cancer. *Lancet* **360**, 683–689 (2002).
40. Schardt, J. A. *et al.* Genomic analysis of single cytokeratin-positive cells from bone marrow reveals early mutational events in breast cancer. *Cancer Cell* **8**, 227–239 (2005).
41. Schmidt-Kittler, O. *et al.* From latent disseminated cells to overt metastasis: genetic analysis of systemic breast cancer progression. *Proc. Natl Acad. Sci. USA* **100**, 7737–7742 (2003).
42. Stoecklein, N. H. *et al.* Direct genetic analysis of single disseminated cancer cells for prediction of outcome and therapy selection in esophageal cancer. *Cancer Cell* **13**, 441–453 (2008).
43. Weckermann, D. *et al.* Perioperative activation of disseminated tumor cells in bone marrow of patients with prostate cancer. *J. Clin. Oncol.* **27**, 1549–1556 (2009).
- References 39 to 43 provide genomic data for single DCCs.**
44. Rasnick, D. Aneuploidy theory explains tumor formation, the absence of immune surveillance, and the failure of chemotherapy. *Cancer Genet. Cytogenet.* **136**, 66–72 (2002).
45. Magbanua, M. J. *et al.* Genomic profiling of isolated circulating tumor cells from metastatic breast cancer patients. *Cancer Res.* **73**, 30–40 (2013).
46. Magbanua, M. J. *et al.* Isolation and genomic analysis of circulating tumor cells from castration resistant metastatic prostate cancer. *BMC Cancer* **12**, 78 (2012).
47. Heitzer, E. *et al.* Complex tumor genomes inferred from single circulating tumor cells by array-CGH and next-generation sequencing. *Cancer Res.* **73**, 2965–2975 (2013).
48. Baudis, M. & Cleary, M. L. Progenetix.net: an online repository for molecular cytogenetic aberration data. *Bioinformatics* **12**, 1228–1229 (2001).
49. Gerlinger, M. *et al.* Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* **366**, 883–892 (2012).
50. Theissen, G. Saltational evolution: hopeful monsters are here to stay. *Theory Biosci.* **128**, 43–51 (2009).
51. Tang, Y. C. & Amon, A. Gene copy-number alterations: a cost-benefit analysis. *Cell* **152**, 394–405 (2013).
52. Schlimok, G. *et al.* Micrometastatic cancer cells in bone marrow: *in vitro* detection with anti-cytokeratin and *in vivo* labeling with anti-17–1A monoclonal antibodies. *Proc. Natl Acad. Sci. USA* **84**, 8672–8676 (1987).
- This paper may be considered as laying the foundation of modern DCC and CTC research by establishing the fundamental concept and the most relevant detection markers.**
53. Cauffman, G., De Rycke, M., Sermon, K., Liebaers, I. & Van de Velde, H. Markers that define stemness in ESC are unable to identify the totipotent cells in human preimplantation embryos. *Hum. Reprod.* **24**, 63–70 (2009).
54. Lu, T. Y. *et al.* Epithelial cell adhesion molecule regulation is associated with the maintenance of the undifferentiated phenotype of human embryonic stem cells. *J. Biol. Chem.* **285**, 8719–8732 (2010).
55. Ng, V. Y., Ang, S. N., Chan, J. X. & Choo, A. B. Characterization of epithelial cell adhesion molecule as a surface marker on undifferentiated human embryonic stem cells. *Stem Cells* **28**, 29–35 (2010).
56. Moll, R., Franke, W. W., Schiller, D. L., Geiger, B. & Krepler, R. The catalog of human cytokeratins: patterns of expression in normal epithelia, tumors and cultured cells. *Cell* **31**, 11–24 (1982).
57. Göttinger, H., Johnson, J. & Riethmüller, G. Biochemical and epitope analysis of the 17–1A membrane antigen. *Hybridoma* **5**, S29–S37 (1986).
58. Gudjonsson, T. *et al.* Isolation, immortalization, and characterization of a human breast epithelial cell line with stem cell properties. *Genes Dev.* **16**, 693–706 (2002).
59. Brabletz, T. *et al.* Variable  $\beta$ -catenin expression in colorectal cancers indicates tumor progression driven by the tumor environment. *Proc. Natl Acad. Sci. USA* **98**, 10356–10361 (2001).
60. Gires, O., Klein, C. A. & Baeuerle, P. A. On the abundance of EpCAM on cancer stem cells. *Nature Rev. Cancer* **9**, 143 (2009).
61. Visvader, J. E. & Lindeman, G. J. Cancer stem cells in solid tumours: accumulating evidence and unresolved questions. *Nature Rev. Cancer* **8**, 755–768 (2008).
62. Riethdorf, S., Wikman, H. & Pantel, K. Biological relevance of disseminated tumor cells in cancer patients. *Int. J. Cancer* **123**, 1991–2006 (2008).
63. Braun, S. *et al.* A pooled analysis of bone marrow micrometastasis in breast cancer. *N. Engl. J. Med.* **353**, 793–802 (2005).
64. Shiozawa, Y. *et al.* Human prostate cancer metastases target the hematopoietic stem cell niche to establish footholds in mouse bone marrow. *J. Clin. Invest.* **121**, 1298–1312 (2011).
- This paper provides the first evidence that DCCs may compete for niches with autochthonous cells.**
65. Engel, J., Emeny, R. T. & Holzel, D. Positive lymph nodes do not metastasize. *Cancer Metastasis Rev.* **31**, 235–246 (2012).
66. Fisher, B. *et al.* Twenty-five-year follow-up of a randomized trial comparing radical mastectomy, total mastectomy, and total mastectomy followed by irradiation. *N. Engl. J. Med.* **347**, 567–575 (2002).
67. Kitchener, H., Swart, A. M., Qian, Q., Amos, C. & Parmar, M. K. Efficacy of systematic pelvic lymphadenectomy in endometrial cancer (MRC ASTEC trial): a randomised study. *Lancet* **373**, 125–136 (2009).
68. Rudenstam, C. M. *et al.* Randomized trial comparing axillary clearance versus no axillary clearance in older patients with breast cancer: first results of International Breast Cancer Study Group. Trial 10–93. *J. Clin. Oncol.* **24**, 337–344 (2006).
69. Veronesi, U., Marubini, E., Mariani, L., Valagussa, P. & Zucali, R. The dissection of internal mammary nodes does not improve the survival of breast cancer patients. 30-year results of a randomised trial. *Eur. J. Cancer* **35**, 1320–1325 (1999).
70. Stoecklein, N. H. & Klein, C. A. Genetic disparity between primary tumours, disseminated tumour cells, and manifest metastasis. *Int. J. Cancer* **126**, 589–598 (2010).
71. Kloosterman, W. P. *et al.* Chromothripsis is a common mechanism driving genomic rearrangements in primary and metastatic colorectal cancer. *Genome Biol.* **12**, R103 (2011).
72. Miranda, E. *et al.* Genetic and epigenetic alterations in primary colorectal cancers and related lymph node and liver metastases. *Cancer* **119**, 266–276 (2013).

73. Shah, S. P. *et al.* Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. *Nature* **461**, 809–813 (2009).
74. Vermaat, J. S. *et al.* Primary colorectal cancers and their subsequent hepatic metastases are genetically different: implications for selection of patients for targeted treatment. *Clin. Cancer Res.* **18**, 688–699 (2012).
75. Campbell, P. J. *et al.* The patterns and dynamics of genomic instability in metastatic pancreatic cancer. *Nature* **467**, 1109–1113 (2010).
76. Stephens, P. J. *et al.* Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell* **144**, 27–40 (2011).
77. Castellari, M. *et al.* Clonal evolution of high-grade serous ovarian carcinoma from primary to recurrent disease. *J. Pathol.* **229**, 515–524 (2013).
78. Ding, L. *et al.* Genome remodelling in a basal-like breast cancer metastasis and xenograft. *Nature* **464**, 999–1005 (2010).
79. Yachida, S. *et al.* Distant metastasis occurs late during the genetic evolution of pancreatic cancer. *Nature* **467**, 1114–1117 (2010).
80. Klein, C. A. Framework models of tumor dormancy from patient-derived observations. *Curr. Opin. Genet. Dev.* **21**, 42–49 (2011).
81. Weedon-Fekjaer, H., Lindqvist, B. H., Vatten, L. J., Aalen, O. O. & Tretli, S. Breast cancer tumor growth estimated through mammography screening data. *Breast Cancer Res.* **10**, R41 (2008).
- The most extensive study on growth rates of breast cancer.**
82. Bos, P. D. *et al.* Genes that mediate breast cancer metastasis to the brain. *Nature* **459**, 1005–1009 (2009).
83. Kang, Y. *et al.* A multigenic program mediating breast cancer metastasis to bone. *Cancer Cell* **3**, 537–549 (2003).
84. Minn, A. J. *et al.* Genes that mediate breast cancer metastasis to lung. *Nature* **436**, 518–524 (2005).
85. Berenblum, I. & Shubik, P. The role of croton oil applications, associated with a single painting of a carcinogen, in tumour induction of the mouse's skin. *Br. J. Cancer* **1**, 379–382 (1947).
86. Van Duuren, B. L., Sivak, A., Katz, C., Seidman, I. & Melchionne, S. The effect of aging and interval between primary and secondary treatment in two-stage carcinogenesis on mouse skin. *Cancer Res.* **35**, 502–505 (1975).
87. Aaronson, S. A. & Todaro, G. J. Basis for the acquisition of malignant potential by mouse cells cultivated *in vitro*. *Science* **162**, 1024–1026 (1968).
88. Lee, L. W., Tsao, M. S., Grisham, J. W. & Smith, G. J. Emergence of neoplastic transformants spontaneously or after exposure to N-methyl-N'-nitro-N-nitrosoguanidine in populations of rat liver epithelial cells cultured under selective and nonselective conditions. *Am. J. Pathol.* **135**, 63–71 (1989).
89. Rubin, H. Cell–cell contact interactions conditionally determine suppression and selection of the neoplastic phenotype. *Proc. Natl Acad. Sci. USA* **105**, 6215–6221 (2008).
90. Johnston, L. A. Competitive interactions between cells: death, growth, and geography. *Science* **324**, 1679–1682 (2009).
91. Levayer, R. & Moreno, E. Mechanisms of cell competition: themes and variations. *J. Cell Biol.* **200**, 689–698 (2013).
92. Moreno, E. Is cell competition relevant to cancer? *Nature Rev. Cancer* **8**, 141–147 (2008).
- References 91 and 92 summarize the evidence for a role for cell competition in cancer.**
93. Mueller, M. M. Inflammation in epithelial skin tumours: old stories and new ideas. *Eur. J. Cancer* **42**, 735–744 (2006).
94. O'Reilly, M. S. *et al.* Angiostatin: a novel angiogenesis inhibitor that mediates the suppression of metastases by a Lewis lung carcinoma. *Cell* **79**, 315–328 (1994).
95. Ruggiero, R. A. *et al.* Tyrosine isomers mediate the classical phenomenon of concomitant tumor resistance. *Cancer Res.* **71**, 7113–7124 (2011).
96. Kaplan, R. N. *et al.* VEGFR1-positive haematopoietic bone marrow progenitors initiate the pre-metastatic niche. *Nature* **438**, 820–827 (2005).
97. McAllister, S. S. *et al.* Systemic endocrine instigation of indolent tumor growth requires osteopontin. *Cell* **133**, 994–1005 (2008).
98. Hiratsuka, S. *et al.* The S100A8-serum amyloid A3–TLR4 paracrine cascade establishes a pre-metastatic phase. *Nature Cell Biol.* **10**, 1349–1355 (2008).
99. Peinado, H. *et al.* Melanoma exosomes educate bone marrow progenitor cells toward a pro-metastatic phenotype through MET. *Nature Med.* **18**, 883–891 (2012).
- Reference 96 to 99 provide insight into how primary tumour-induced systemic alterations are linked to metastatic success of DCCs.**
100. Ghajar, C. M. *et al.* The perivascular niche regulates breast tumour dormancy. *Nature Cell Biol.* **15**, 807–817 (2013).
101. Klein, C. A. & Stoecklein, N. H. Lessons from an aggressive cancer: evolutionary dynamics in esophageal carcinoma. *Cancer Res.* **69**, 5285–5288 (2009).
102. Tang, Y. C., Williams, B. R., Siegel, J. J. & Amon, A. Identification of aneuploidy-selective antiproliferation compounds. *Cell* **144**, 499–512 (2011).
103. Steeg, P. S. The right trials. *Nature* **485**, S58–S59 (2012).
104. Hölzel, D., Eckel, R., Emeny, R. T. & Engel, J. Distant metastases do not metastasize. *Cancer Metastasis Rev.* **29**, 737–750 (2010).

**Acknowledgements** I am indebted to T. Perry for his critical reading of the manuscript and his invaluable suggestions to improve it. I am also grateful to S. Pausch for her help with the figures.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The author declares no competing financial interests. Readers are welcome to comment on the online version of this article at [go.nature.com/a5mpwv](http://go.nature.com/a5mpwv). Correspondence should be addressed to C. K. ([christoph.klein@ukr.de](mailto:christoph.klein@ukr.de)).

# Cerebral organoids model human brain development and microcephaly

Madeline A. Lancaster<sup>1</sup>, Magdalena Renner<sup>1</sup>, Carol-Anne Martin<sup>2</sup>, Daniel Wenzel<sup>1</sup>, Louise S. Bicknell<sup>2</sup>, Matthew E. Hurles<sup>3</sup>, Tessa Homfray<sup>4</sup>, Josef M. Penninger<sup>1</sup>, Andrew P. Jackson<sup>2</sup> & Juergen A. Knoblich<sup>1</sup>

**The complexity of the human brain has made it difficult to study many brain disorders in model organisms, highlighting the need for an *in vitro* model of human brain development. Here we have developed a human pluripotent stem cell-derived three-dimensional organoid culture system, termed cerebral organoids, that develop various discrete, although interdependent, brain regions. These include a cerebral cortex containing progenitor populations that organize and produce mature cortical neuron subtypes. Furthermore, cerebral organoids are shown to recapitulate features of human cortical development, namely characteristic progenitor zone organization with abundant outer radial glial stem cells. Finally, we use RNA interference and patient-specific induced pluripotent stem cells to model microcephaly, a disorder that has been difficult to recapitulate in mice. We demonstrate premature neuronal differentiation in patient organoids, a defect that could help to explain the disease phenotype. Together, these data show that three-dimensional organoids can recapitulate development and disease even in this most complex human tissue.**

Mammalian brain development begins with the expansion of the neuroepithelium to generate radial glial stem cells (RGs)<sup>1</sup>. These RGs divide at the apical surface within the ventricular zone (VZ) to generate neurons and intermediate progenitors. Intermediate progenitors populate the adjacent subventricular zone (SVZ), whereas neurons migrate through the intermediate zone to populate specific layers within the cortical plate. In humans, the organization of progenitor zones is markedly more elaborate; the SVZ is split by an inner fibre layer (IFL) into an inner SVZ and an outer SVZ (OSVZ)<sup>2</sup>. The OSVZ represents an entirely separate progenitor layer populated by intermediate progenitors and a unique stem cell subset termed outer radial glia (oRG)<sup>3,4</sup>, which are only present to a limited degree in rodents<sup>5</sup>. Both the IFL and OSVZ are completely absent in mice<sup>6</sup>. These key differences allow for the striking expansion in neuronal output and brain size seen in humans<sup>7,8</sup>.

Primary microcephaly is a neurodevelopmental disorder in which brain size is markedly reduced<sup>9</sup>. Autosomal-recessive mutations have been identified in several genes, all of which encode proteins localizing to the mitotic spindle apparatus<sup>10</sup>. Heretofore, primary microcephaly pathogenesis has primarily been examined in mouse models. However, mouse mutants for several of the known genes<sup>11–14</sup> have failed to recapitulate the severely reduced brain size seen in human patients.

Given the dramatic differences between mice and humans, methods that recapitulate paradigms of human brain development *in vitro* have enormous potential. Although considerable progress has been made for *in vitro* models of whole-organ development for other systems—such as intestine<sup>15</sup>, pituitary<sup>16</sup> and retina<sup>17,18</sup>—a three-dimensional culture model of the developing brain as a whole has not been established. Previous studies have modelled certain isolated neural tissues *in vitro*<sup>19–23</sup>, including dorsal cerebral cortical tissue, which could recapitulate many aspects of early development. However, later events such as the formation of discrete cortical layers with stereotypical inside-out organization as well as human characteristics, such as the presence of oRG cells and the unique organization of progenitor zones, were not described.

Here we describe a three-dimensional culture system for deriving brain tissue *in vitro*. These so-called cerebral organoids develop a variety

of regional identities organized as discrete domains capable of influencing one another. Furthermore, cerebral cortical regions display an organization similar to the developing human brain at early stages, as well as the presence of a considerable oRG population. Moreover, cerebral cortical neurons mature to form various pyramidal identities with modest spatial separation. Finally, we use patient-derived induced pluripotent stem (iPS) cells and short hairpin RNA (shRNA) in these organoids to model CDK5RAP2-dependent pathogenesis of microcephaly, a disorder that has been difficult to model in mice.

## Generation of cerebral organoids

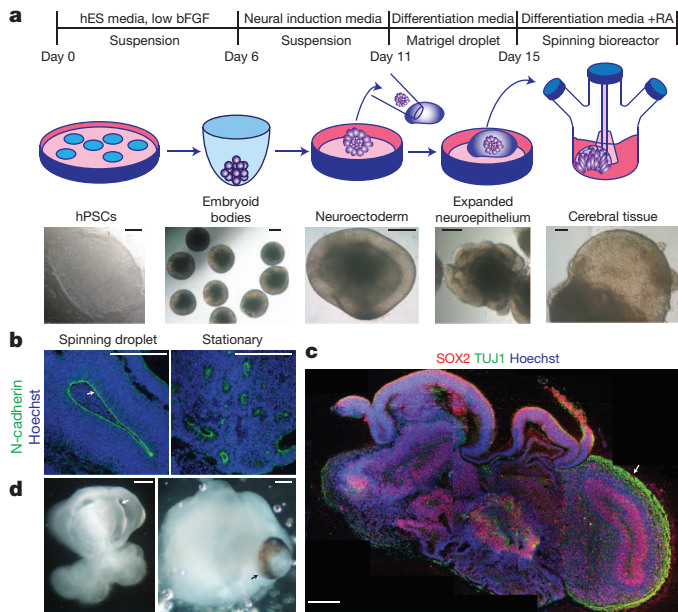
Recent progress with *in vitro* models of various organ systems has demonstrated the enormous self-organizing capacity for pluripotent stem cells to form whole tissues<sup>15,18</sup>. We built upon this concept and developed a protocol without the use of patterning growth factors, focusing instead on improving growth conditions and providing the environment necessary for intrinsic cues to influence development. We began with a modified approach to generate neuroectoderm from embryoid bodies<sup>24</sup>. Neuroectodermal tissues were then maintained in three-dimensional culture and embedded in droplets of Matrigel to provide a scaffold for more complex tissue growth. These Matrigel droplets were then transferred to a spinning bioreactor to enhance nutrient absorption (Fig. 1a). This method led to rapid development of brain tissues, which we termed cerebral organoids, requiring only 8–10 days for the appearance of neural identity and 20–30 days for defined brain regions to form.

Cerebral organoids at early stages (15–20 days) formed large, continuous neuroepithelia surrounding a fluid-filled cavity reminiscent of a ventricle with characteristic apical localization of the neural specific N-cadherin (Fig. 1b). Furthermore, the neuroepithelium was larger and more continuous than tissues generated using a stationary approach<sup>20</sup>, which instead formed an aggregate of several small rosette-like neuroepithelia (Fig. 1b and Extended Data Fig. 1a).

Although tissues reached maximal size by 2 months, cerebral organoids formed large (up to 4 mm in diameter), complex heterogeneous

<sup>1</sup>Institute of Molecular Biotechnology of the Austrian Academy of Science (IMBA), Vienna 1030, Austria. <sup>2</sup>MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh EH4 2XU, UK. <sup>3</sup>Wellcome Trust Sanger Institute, Cambridge CB10 1SA, UK. <sup>4</sup>Department of Clinical Genetics, St. George's University, London SW17 0RE, UK.





**Figure 1 | Description of cerebral organoid culture system.** **a**, Schematic of the culture system described in detail in Methods. Example images of each stage are shown. bFGF, basic fibroblast growth factor; hES, human embryonic stem cell; hPSCs, human pluripotent stem cells; RA, retinoic acid. **b**, Neuroepithelial tissues generated using this approach (left) exhibited large fluid-filled cavities and typical apical localization of the neural N-cadherin (arrow). These tissues were larger and more continuous than tissues grown in stationary suspension without Matrigel (right). **c**, Sectioning and immunohistochemistry revealed complex morphology with heterogeneous regions containing neural progenitors (SOX2, red) and neurons (TUJ1, green) (arrow). **d**, Low-magnification bright-field images revealing fluid-filled cavities reminiscent of ventricles (white arrow) and retina tissue, as indicated by retinal pigmented epithelium (black arrow). Scale bars, 200  $\mu$ m.

tissues, which could survive indefinitely (currently up to 10 months) when maintained in a spinning bioreactor. Histological and gross morphological analysis revealed regions reminiscent of cerebral cortex, choroid plexus, retina and meninges (Fig. 1c, d and Extended Data Fig. 1b). Notably, tissues typically reached a size limit, probably because of the lack of a circulatory system and limitations in oxygen and nutrient exchange. Consistent with this, extensive cell death was visible in the core of these tissues (Extended Data Fig. 1c), whereas the various brain regions developed along the exterior. Furthermore, cerebral organoids could be reproducibly generated with similar overall morphology and complexity from both human embryonic stem (ES) cells and iPS cells (Extended Data Fig. 1d, e).

### Cerebral organoids display discrete brain regions

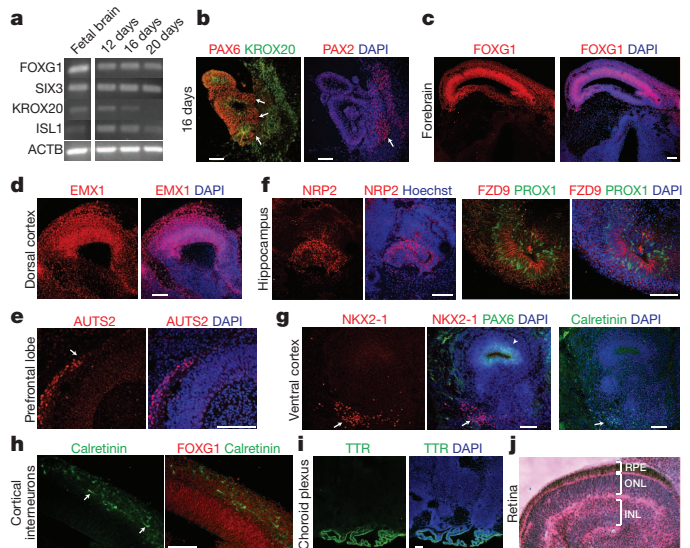
Brain development *in vivo* exhibits a striking degree of heterogeneity and regionalization as well as interdependency of various brain regions. Histological analysis suggested that human cerebral organoids might similarly display heterogeneous brain regions. To examine this further, we first tested the efficiency of initial neural induction in these tissues by performing reverse transcriptase PCR (RT-PCR) for several markers of pluripotency and neural identity (Extended Data Fig. 2a). As expected, pluripotency markers OCT4 (also known as POU5F1) and NANOG diminished during the course of organoid differentiation, whereas neural identity markers SOX1 and PAX6 were upregulated, indicating successful neural induction.

To test for early brain regionalization in whole organoids, we performed RT-PCR for forebrain (FOXP1 and SIX3) and hindbrain (KROX20 (also known as EGR2) and ISL1) markers (Fig. 2a), revealing the presence of both populations within the tissue. However, as tissue development proceeded, forebrain markers remained highly expressed whereas hindbrain markers decreased, reminiscent of the

developmental expansion of forebrain tissue during human brain development<sup>25</sup>.

In order to test whether cells with these brain region identities developed as discrete regions within the organoids, as gross morphology would suggest, or were randomly interspersed within the tissue, we performed immunohistochemical staining for markers of forebrain, midbrain and hindbrain identities during early development of these tissues (16 days; Fig. 2b and Extended Data Fig. 2b). PAX6 expression revealed several regions of forebrain identity, and OTX1 and OTX2 expression marked forebrain/midbrain identity. These regions were located adjacent to regions that lacked these markers but that were positive for hindbrain markers GBX2, KROX20 and PAX2, which was reminiscent of the early mid–hindbrain boundary, suggesting similar regional communication and probably mutual repression.

*In vivo* brain development involves increasing refinement of regional specification. Therefore, we examined further-developed cerebral organoid tissues for regional subspecification. We performed staining for the forebrain marker FOXG1 (Fig. 2c), which labelled regions displaying typical cerebral cortical morphology. Many of these regions were also positive for EMX1 (Fig. 2d), indicating dorsal cortical identity. We also tested for further subregionalization by staining for cortical lobe markers, namely AUTS2, a marker of prefrontal cortex<sup>26</sup> (Fig. 2e); TSHZ2, a marker of the occipital lobe<sup>26</sup> (Extended Data Fig. 2c); and LMO4, a marker of frontal and occipital lobes but absent in parietal lobes<sup>26</sup> (Extended Data Fig. 2c). These markers could be seen in neurons labelling distinct regions of dorsal cortex, suggesting subspecification of cortical lobes.



**Figure 2 | Human cerebral organoids recapitulate various brain region identities.** **a**, RT-PCR for forebrain markers (FOXP1 and SIX3) and hindbrain markers (KROX20 and ISL1) at 12, 16 and 20 days of differentiation. Human fetal brain complementary DNA was used as positive control.

**b**, Immunohistochemistry in serial sections for the forebrain marker PAX6 (red, left) and the hindbrain markers KROX20 (green, left) and PAX2 (red, right) at 16 days of differentiation. Note the juxtaposition reminiscent of the mid–hindbrain boundary (arrows). DAPI (4',6-diamidino-2-phenylindole) marks nuclei (blue). **c–i**, Staining for various brain region identities: forebrain, FOXG1 (c); dorsal cortex, EMX1 (d); prefrontal cortex (note the discrete boundary, arrow), AUTS2 (e); hippocampus, NRP2, FZD9, PROX1 (f); ventral forebrain, NKX2-1 (g) and choroid plexus, TTR (i). **g**, Staining for adjacent ventral (arrow) and dorsal (PAX6, arrowhead) forebrain and for calretinin (green) in a serial section revealing cortical interneurons in the ventral region (arrow). Calretinin interneurons within dorsal cortex (h) exhibit typical morphology of tangential migration (arrows). **j**, Haematoxylin and eosin staining of retinal tissue exhibiting stereotypical layering: retinal pigment epithelium (RPE), outer nuclear layer (ONL) and inner nuclear layer (INL). Scale bars, 100  $\mu$ m.

Furthermore, staining for markers of the hippocampus (Fig. 2f) and ventral forebrain (Fig. 2g) revealed specification of these regions, although they did not organize to form the overall structure seen *in vivo*. Notably, interneurons produced in ventral forebrain regions exhibited a morphology and location consistent with migration from ventral to dorsal tissues (Extended Data Fig. 2d). Within the dorsal cortex, these neurons displayed neurites parallel to the apical surface, reminiscent of the migratory extensions seen in tangential migration *in vivo* (Fig. 2h). Notably, calretinin<sup>+</sup> interneurons were absent from the dorsal cortex of organoids lacking a ventral region (4 out of 4 NKX2-1-negative organoids), suggesting that interneurons originate in the ventral forebrain to migrate to the dorsal cortex. This suggests that distant regions can influence one another in developing cerebral organoids.

Finally, other brain structures could be observed, namely choroid plexus (Fig. 2i) and even immature retina (Fig. 2j). Overall, all tissues examined displayed regions with dorsal cortical morphology (35 out of 35; 100%), most displayed choroid plexus (25 out of 35; 71%) and several displayed ventral forebrain identity as determined by NKX2-1 immunoreactivity (12 out of 35; 34%), whereas only a few displayed retinal tissue (determined by presence of retinal pigmented epithelium; 4 out of 35; 11%). These results indicate that cerebral organoids developed a variety of brain region identities organized into discrete, although interdependent, domains.

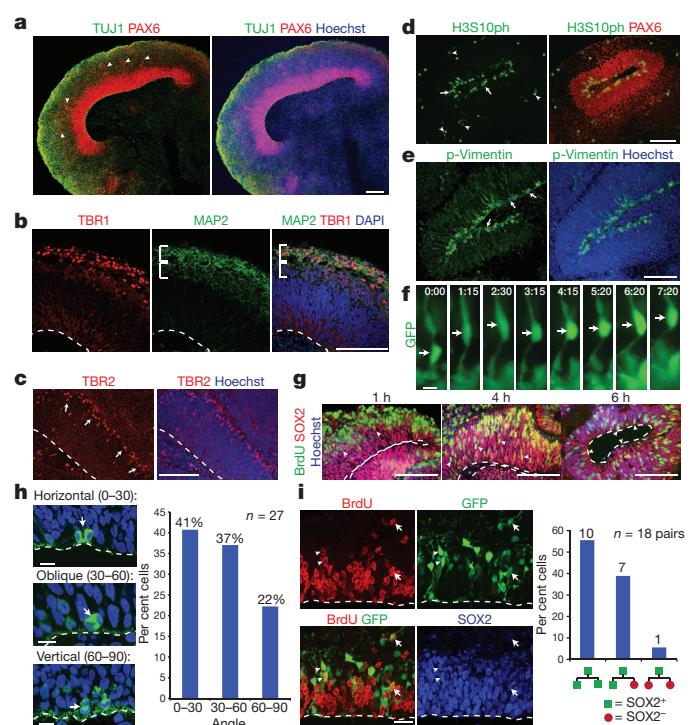
### Recapitulation of dorsal cortical organization

The most dramatic changes in brain evolution from rodent to human affect the dorsal cortex. Therefore, we analysed the organization of dorsal cortical regions within cerebral organoids. Staining for markers of RGs and newborn neurons (Fig. 3a) revealed typical organization into a layer reminiscent of the VZ with neurons located at the basal surface. Staining for TBR1 (Fig. 3b) revealed proper development of neural identity and radial migration to the developing preplate (precursor to the cortical plate). Furthermore, staining for neural progenitor and neural-specific BAF (mammalian SWI/SNF chromatin-remodelling complex) components revealed the characteristic switch in chromatin-remodelling complexes during neural fate specification<sup>27,28</sup> (Extended Data Fig. 3a). Finally, staining for the intermediate progenitor marker TBR2 (also known as EOMES) (Fig. 3c) revealed the presence of intermediate progenitors adjacent to the VZ. Thus, dorsal cortical tissues displayed typical progenitor zone organization.

In both mice and humans, cortical progenitors undergo a stereotypical nuclear movement called interkinetic nuclear migration (IKNM). Mitotic divisions occur at the apical surface of the VZ whereas the nuclei of cells in S phase are located on the basal side of the VZ<sup>29</sup>. We stained for the mitotic marker phospho-histone H3 (Fig. 3d) and observed most of the cells dividing at the apical surface. Similar observations were evident when we stained for phospho-vimentin (p-vimentin, Fig. 3e), a marker of mitotic RGs. In addition, as this marker labels the entire cell, we could observe basal cellular processes typical of RGs, which extended to the outer surface of these tissues (Extended Data Fig. 3b). Thus, RGs of cerebral organoids exhibited typical behaviour and morphological features.

To examine this in more detail, we used a method to label individual RGs for morphology and live imaging experiments. In the mouse brain, individual cells can be labelled by *in utero* electroporation of fluorescent-protein-expressing plasmids. Similarly, we injected green fluorescent protein (GFP) plasmid into fluid-filled cavities of cerebral organoids and electroporated RGs adjacent to these ventricle-like cavities (Extended Data Fig. 3c). This approach led to reproducible expression of GFP in RGs, revealing typical morphology at various stages of development: early pseudostratified neuroepithelium (Extended Data Fig. 3d) followed by later bipolar morphology with extended apical and basal processes (Extended Data Fig. 3e, f).

To test for IKNM, we performed live imaging of GFP-electroporated RGs in cerebral organoids and observed many examples of RGs that displayed movement of the cell body along the apical and basal processes



**Figure 3 | Stereotypical organization and behaviour of progenitors.**

**a**, Immunohistochemistry for neurons (TUJ1, green) and RGs (PAX6, red) in a large dorsal cortical region. Note the additional PAX6<sup>+</sup> RGs located outside the VZ (arrowheads), reminiscent of oRGs. **b**, Staining for the preplate marker TBR1 (red) and neuronal marker MAP2 (green) revealing superficial preplate (upper bracket) and underlying neuronal intermediate zone-like layer (lower bracket). **c**, Staining for the intermediate progenitor marker TBR2 (red) revealing SVZ localization of intermediate progenitors (arrows). **d**, Staining for phospho-histone H3 (H3S10ph green) to mark RG (PAX6<sup>+</sup>) in mitosis. Arrows mark apical surface divisions; arrowheads mark SVZ divisions. **e**, p-Vimentin (green) staining for mitotic RG, which primarily divide at the apical surface (arrows). **f**, Frames from live imaging of GFP-electroporated RGs showing cell body movement (arrows). Time shown in h:min. **g**, BrdU pulse-chase revealing progressive IKNM of BrdU-labelled nuclei (green, arrowheads) from basal VZ (1 h) to a more apical position (4–6 h). **h**, Quantification of RG division orientation displayed in bins of 0–30 (horizontal), 30–60 (oblique) and 60–90 (vertical) degrees. **i**, Lineage tracing in GFP-electroporated and BrdU-pulsed tissues to mark daughter cell pairs following 16-h chase revealing symmetric (arrowheads) and asymmetric (arrows) fates indicated by SOX2 staining (blue). Quantification for 18 cell pairs from three cortical tissues. Numbers above bars are absolute cell numbers. Dashed line indicates apical surface (**b**, **c**, **g**, **h**, **i**). Scale bars, 100  $\mu$ m (**a–e**, **g**), 10  $\mu$ m (**f**, **h**) and 20  $\mu$ m (**i**).

(Fig. 3f and Supplementary Video 1). Furthermore, we performed pulse-chase experiments with the S-phase marker BrdU (Fig. 3g) and could observe a shift in RG nuclei from outer VZ localization towards the apical surface with time (Fig. 3g).

RGs in the VZ of rodents exhibit biased spindle orientation, predominantly horizontal, parallel to the ventricular surface<sup>30–33</sup>. To examine whether RGs in human cerebral organoids exhibited a similar orientation bias, we used p-vimentin staining to examine the plane of division in mitotic RGs (Extended Data Fig. 3g). We observed mainly horizontal orientations (41%) (Fig. 3h), somewhat similar to the orientation bias observed in other mammals. However, we also observed abundant oblique (37%) and vertical (22%) orientations, which were more abundant in these human tissues than has been described for rodent neocortex<sup>30,31,34,35</sup>. Interestingly, these measurements reflected the same trend recently described in the human brain<sup>36</sup>, suggesting that the cerebral organoids could recapitulate aspects of human cortical development.

We examined further the fate potential of these divisions to test whether organoid RGs could divide symmetrically and/or asymmetrically. We



performed electroporation of GFP followed by a short 1-h BrdU pulse and a 16-h chase to lineage trace divisions of a small minority of cells. We examined double-labelled daughter-cell pairs and could observe both symmetric proliferative fate outcomes, as well as asymmetric outcomes (Fig. 3i). This suggests that the RGs in these human tissues can undergo both symmetric and asymmetric divisions.

### Formation of functional cerebral cortical neurons

The formation of the radially organized cortical plate begins with the formation of its precursor, the preplate. To test for this initial organization, we stained 30-day organoids for TBR1, a marker of the preplate<sup>37</sup>, as well as MAP2, a neuronal marker<sup>38</sup> (Fig. 4a). This revealed the presence of a basal neural layer reminiscent of the preplate, and an apically adjacent region reminiscent of the intermediate zone. Furthermore, we could observe reelin<sup>+</sup> neurons along the basal surface, suggesting the presence of Cajal–Retzius cells, a cell population important in the generation of cortical plate architecture<sup>39</sup>.

The stereotypical layered structure of the mammalian cortical plate is generated inside-out with later-born neurons migrating through existing layers to populate more superficial layers<sup>40</sup>. Although previous methods of deriving cortical neurons have been able to generate distinct layer identities<sup>20,23,41</sup>, they have been unable to recapitulate this spatial separation. To test whether this organization could be recapitulated in cerebral organoids, we stained for cortical layer markers. In less-developed tissues (30 day), early-born CTIP2 (also known as BCL11B)<sup>+</sup> neurons were located adjacent and internal to the TBR1<sup>+</sup> preplate, suggesting initiation of cortical plate layer formation (Extended Data Fig. 4a). Furthermore, neurons exhibited rudimentary separation into an early-born deep layer (CTIP2<sup>+</sup>) and a late-born superficial layer (SATB2<sup>+</sup> and BRN2 (also known as POU3F2)<sup>+</sup>) (Extended Data Fig. 4b), which became more distinct as tissues developed (75 days) (Fig. 4c). Finally, a cell-poor region reminiscent of the marginal zone was also evident (Extended Data Fig. 1b). Notably, although this modest spatial separation was an improvement upon other *in vitro* methods, organoids could not recapitulate the same degree of mature layer organization as *in vivo*, suggesting that further cues are needed to generate the complex stereotypical layer II–VI organization.

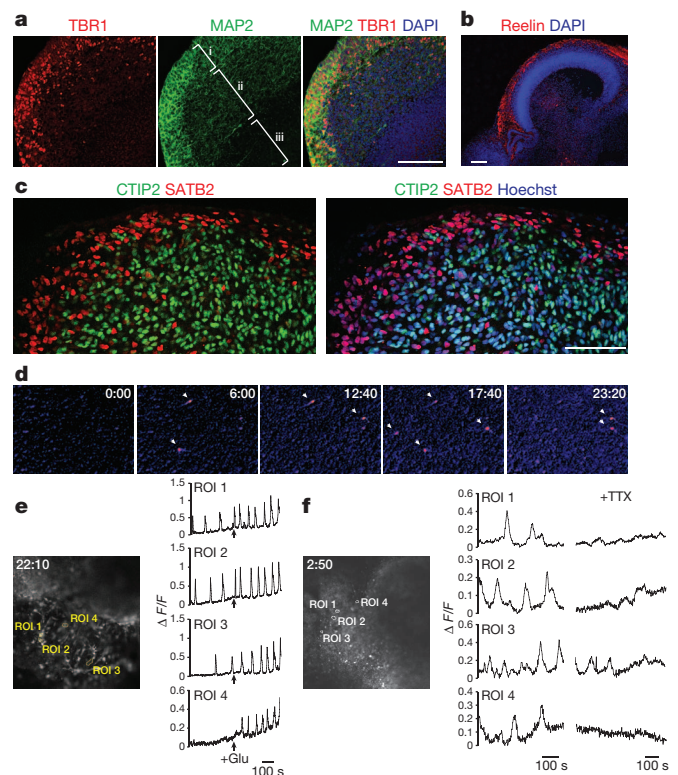
*In vivo*, dorsal cortical neurons mature and extend long-range axons<sup>42</sup>. To test for these characteristics, we performed GFP electroporation and examined neuronal morphology. GFP-labelled axon projections displayed complex branching and growth cone behaviour (Extended Data Fig. 4c) and projected long-range axons in a manner reminiscent of axon bundling (Extended Data Fig. 4d)<sup>43</sup>.

Finally, we tested whether neurons within cerebral organoids could exhibit neural activity by performing calcium dye imaging to detect Ca<sup>2+</sup> oscillations<sup>44</sup>, which revealed spontaneous calcium surges in individual cells (Fig. 4d, Extended Data Fig. 4e and Supplementary Videos 2, 3). Furthermore, we applied exogenous glutamate (Fig. 4e) and observed more frequent calcium spikes, indicating glutamatergic receptor activity. Finally, we performed action potential blockade by application of tetrodotoxin and observed dampened calcium surges, indicating that calcium spikes were dependent upon neuronal activity (Fig. 4f).

### Recapitulation of human cortical organization

Human brain development exhibits unique progenitor zone organization<sup>2,7</sup>. To test whether features of human brain development were recapitulated in cerebral organoids, we examined the distribution of SOX2<sup>+</sup> progenitors and observed a population displaced from the apical surface (Fig. 5a and Extended Data Fig. 5a), consistent with an oRG identity. Furthermore, these fairly abundant oRGs appeared separated from the apical VZ by a TUJ1<sup>+</sup> fibre layer (Fig. 5a) reminiscent of the IFL. This organization suggests that human cerebral organoids could recapitulate at least some aspects of human cortical development.

To rule out the possibility that this OSVZ-like organization was an *in vitro* artefact, we adapted the method to mouse ES cells to generate



**Figure 4 | Organization and maturation of cerebral cortical neurons.**

**a**, Immunohistochemical staining at day 30 showing preplate (TBR1) with early signs of radial organization (MAP2, bracket i) and the presence of an intermediate zone-like layer (bracket ii) adjacent to the VZ/SVZ (bracket iii). DAPI marks nuclei (blue). **b**, Reelin staining indicating Cajal–Retzius cells along the basal surface of dorsal cortical tissue. **c**, Staining for early-born (CTIP2) and late-born (SATB2) neurons at 75 days differentiation reveals separation and rudimentary inside-out organization. **d**, False-colour heat map frames from Fluo-4 calcium live imaging revealing spontaneous calcium surges (arrowheads). Time is displayed in min:sec. **e**, Single-cell tracings of calcium surges with glutamate application (regions of interest (ROI) outlined in left panel) as measured by change in fluorescence (arbitrary units). Arrows mark the time of addition of glutamate. **f**, Single-cell tracing (ROIs marked in image at left) of calcium surges before (left) and after (right) the addition of tetrodotoxin (TTX). Scale bars, 100  $\mu$ m.

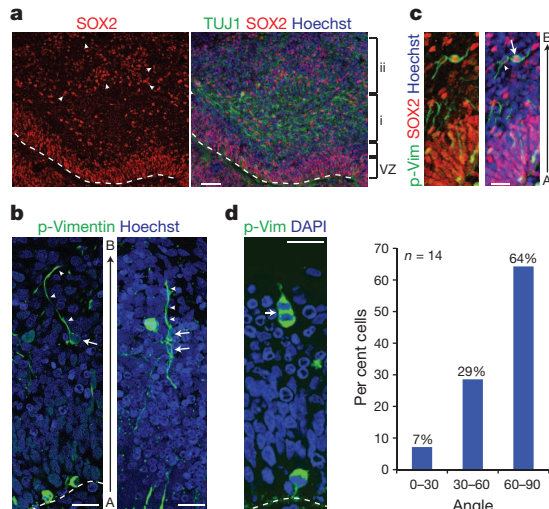
mouse cerebral organoids and examined whether a similar organization was present (Extended Data Fig. 5b, c). We observed much smaller cortical tissues in mouse organoids compared with humans, and only occasional oRGs that did not accumulate in an OSVZ-like region. These results indicate that OSVZ- and IFL-like layers are specific to human organoids.

We performed p-vimentin staining to examine the morphology of human oRGs, which revealed basal processes but a lack of apical processes (Fig. 5b), a hallmark of oRGs<sup>3,4</sup>. This suggests that these basally displaced SOX2<sup>+</sup> and p-vimentin<sup>+</sup> progenitors indeed represent human oRGs. We also examined the division mode of oRGs and could identify daughter cell pairs in which only one daughter cell maintained SOX2 expression (Fig. 5c), suggesting asymmetric division. Furthermore, we measured spindle orientation and found that nearly all oRGs divided vertically (Fig. 5d). This division mode is remarkably similar to the findings recently described in human brain slice cultures<sup>36</sup>, suggesting recapitulation of human oRG behaviour in cerebral organoids.

### Cerebral organoids model human microcephaly

Because disorders affecting human brain development have often proved difficult to recapitulate in animal models, we tested whether organoids could be used to model neurodevelopmental disorders. We identified a patient with severe microcephaly (−13.2 s.d. below mean for age and sex, Fig. 6a) and reduced stature (−6.7 s.d., Supplementary Text and





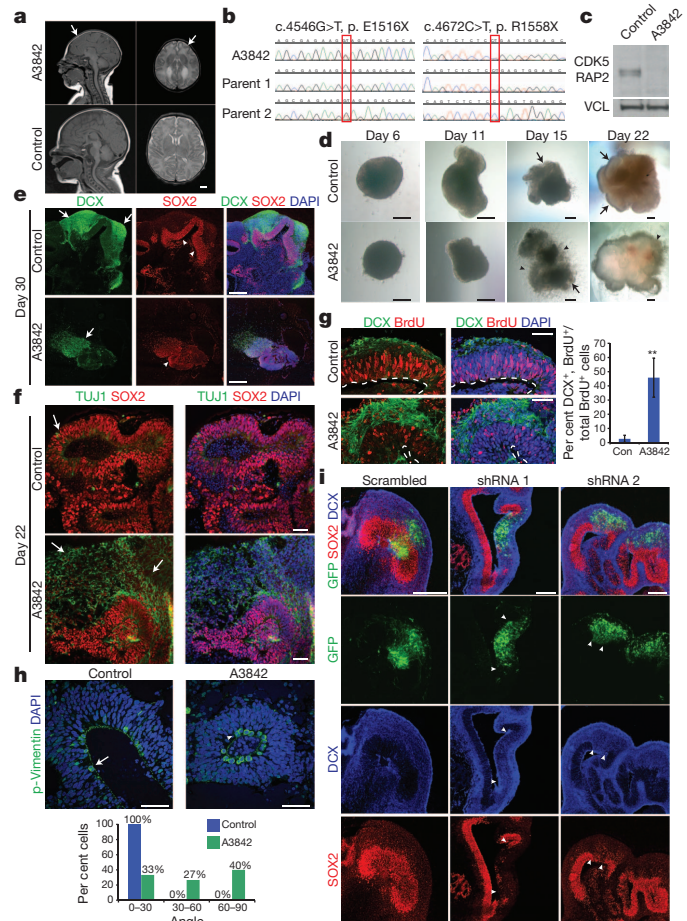
**Figure 5 | Cerebral organoids produce oRGs with typical morphology and behaviour.** **a**, Immunohistochemical staining for RGs (SOX2) and neuronal processes (TUJ1) reveals the presence of oRGs (arrowheads) organized similar to human cortical development (OSVZ-like layer, bracket ii) separated from the VZ by a layer of TUJ1<sup>+</sup> fibres similar to IFL (bracket i). **b**, Staining for p-vimentin revealing dividing oRGs (arrows) with typical morphology: possessing a basal process (arrowheads) but lacking an apical process. Right, a daughter cell pair showing unequal inheritance of the basal process. Apical–basal orientation indicated by A→B. **c**, Staining for p-vimentin in a recently divided daughter cell pair with asymmetric fates: one oRG (arrow, SOX2<sup>+</sup>), whereas the other lacks SOX2 expression (arrowhead). **d**, Orientation of division of a mitotic oRG in anaphase revealing vertical (60–90 degrees) orientation relative to the apical surface (dashed line). Quantification of spindle orientation for 14 anaphase oRGs from 6 different cortical tissues. Scale bars, 50  $\mu$ m (**a**), 20  $\mu$ m (**b–d**).

Extended Data Fig. 6a), who, as determined through exome sequencing and confirmed by capillary sequencing (Fig. 6b), had compound heterozygous truncating mutations in *CDK5RAP2*. Skin fibroblasts from this patient revealed a loss of the CDK5RAP2 protein (Fig. 6c and Extended Data Fig. 6b) suggesting loss of function, consistent with previously reported *CDK5RAP2* mutations in primary microcephaly patients<sup>45,46</sup>.

We performed reprogramming of patient skin fibroblasts using lentiviral delivery of the four well-described reprogramming factors: OCT4, SOX2, MYC and KLF4 (refs 47, 48). We generated many independent clones of iPS cells and characterized four of these for morphology and pluripotency. All four lines exhibited similar doubling times and colony morphology that was indistinguishable from control iPS cells (Extended Data Fig. 7a). All lines could form embryoid bodies and stained positive for the pluripotency marker alkaline phosphatase (Extended Data Fig. 7b).

We performed cerebral organoid culture from all four lines and observed smaller embryoid bodies, which when transferred to neural induction failed to develop further (Extended Data Fig. 7c). We proposed that because the patient also displayed reduced stature, perhaps overall embryoid body growth was perturbed. Therefore, in order to allow the embryoid bodies to develop to a comparable size as control, we modified the protocol slightly by increasing the initial iPS cell number. This modification allowed for the formation of neuroectoderm and subsequent neural tissue for analysis, which revealed smaller neuroepithelial tissues and a large degree of neuronal outgrowth compared with control tissues (Fig. 6d and Extended Data Fig. 7d). Immunohistochemical staining for progenitors and neurons revealed smaller neural tissues with only very few progenitor regions (Fig. 6e). These overall smaller neural tissues were reminiscent of the reduced brain size seen in the patient.

These patient-derived cerebral organoids provided a unique opportunity to examine the cause of the hypoplasia seen in microcephaly. We therefore examined an earlier stage (22 days) and observed that, whereas control tissues displayed abundant, large neuroepithelial tissues composed



**Figure 6 | Cerebral organoid modelling of microcephaly.** **a**, Magnetic resonance imaging scan from patient A3842 taken at birth (top) compared with age-matched control (bottom) showing brain and head size reduction and simplified cortical folding (arrows). Sagittal T1 (left) and axial T2 (right) images. Scale bar, 1 cm. **b**, Sequencing chromatograms demonstrating compound heterozygous nonsense mutations inherited from each parent. **c**, CDK5RAP2 protein is undetectable on immunoblotting of patient cell lysate (A3842) compared with control skin fibroblasts. Vinculin (VCL), loading control. **d**, Representative bright-field images of control and patient-derived cerebral organoids (A3842 line 1M, all lines shown in Extended Data Fig. 7d) at 6, 11, 15 and 22 days of differentiation. Control exhibits large fluid-filled cortical regions (arrows), whereas patient-derived exhibits increased outgrowth (arrowheads). **e**, Immunohistochemistry in control and patient-derived (10M) tissues at day 30 of differentiation revealing fewer neurons (DCX, arrows) and smaller progenitor zones (SOX2, arrowheads). **f**, Staining at day 22 showing increased neurons (TUJ1, arrows) in patient-derived tissue (14B). **g**, BrdU pulse-chase in control and patient-derived organoids (14B) showing higher percentage of BrdU<sup>+</sup> cells with neural identity and less in the VZ compared with control. Results quantified on right. Error bars are s.d. **\*\*P** < 0.01, Student's *t*-test. *n* = 3 organoids for each condition (300 cells total for control, 204 cells for patient). **h**, p-Vimentin staining in control and patient-derived tissues (14B) showing RG mitotic divisions. Control RGs at anaphase divided exclusively horizontal (0–30 degree angle, arrow) whereas patient RGs displayed many oblique and vertical orientations (arrowhead). Results quantified on right (*P* < 0.01, 2 × 3 Fisher's exact test, *n* = 11 cells for control, *n* = 15 cells for patient-derived, from >5 cortical regions each). **i**, Human ES cell organoids co-electroporated with GFP and scrambled or *CDK5RAP2* shRNAs and examined after 5 days. Electroporated regions (demarcated by arrowheads) exhibit loss of SOX2<sup>+</sup> progenitors and increased DCX<sup>+</sup> neurons. Scale bars, 200  $\mu$ m (**d**, **e**, **i**), 50  $\mu$ m (**f–h**).

of progenitors, patient-derived tissues displayed only occasional neuroepithelial regions (Extended Data Fig. 7e). Furthermore, these tissues displayed decreased RGs and increased neurons compared with control (Fig. 6f and Extended Data Fig. 7f), suggesting premature neural

differentiation. To test this possibility, we performed BrdU pulse-chase experiments (Fig. 6g), which revealed a marked increase in the number of BrdU<sup>+</sup>/doublecortin (DCX)<sup>+</sup> cells in patient organoids, consistent with premature neurogenic non-proliferative divisions.

We next examined radial glial spindle orientation and observed that, whereas control tissues at this early stage (22 days) displayed exclusively horizontal orientations (Fig. 6h), patient organoids displayed many oblique and vertical oriented spindles (Fig. 6h). These results could explain the patient tissue phenotype because precise horizontal orientation of the spindle is necessary for early symmetric expansion of neural stem cells<sup>32</sup>.

Finally, we tested whether the phenotype could be rescued by reintroducing CDK5RAP2 protein. We performed co-electroporation of GFP and CDK5RAP2 into day-12 patient organoids and examined 6 days later. Because high overexpression of CDK5RAP2 was toxic (data not shown), the cells with high GFP signal did not survive to this time point. However, we could observe regions in CDK5RAP2-electroporated tissues with larger neuroepithelium compared with tissues electroporated only with GFP (Extended Data Fig. 7g). This effect could be due to surviving cells with a low level of CDK5RAP2 re-expression. Supporting this interpretation, staining for GFP (Extended Data Fig. 7h) revealed many low-level GFP<sup>+</sup> cells in CDK5RAP2-co-electroporated patient organoids with radial glial morphology (54% ± 2 s.e.m.; *n* = 74 cells from three tissues). By contrast, GFP<sup>+</sup> cells in patient organoids electroporated with GFP alone exhibited mainly neuronal morphology with significantly fewer RG (19% ± 11 s.e.m.; *n* = 102 cells from three tissues; *P* < 0.05, Student's *t*-test). Thus, we conclude that the phenotype is specific to the loss of CDK5RAP2.

As a further independent approach, we performed RNAi knock-down of *CDK5RAP2* by co-electroporating GFP with two independent shRNAs found to knock down endogenous *CDK5RAP2* (Extended Data Fig. 8a). Both shRNAs led to a considerable loss of SOX2<sup>+</sup> progenitors and an increase in DCX<sup>+</sup> neurons (Fig. 6i and Extended Data Fig. 8b), reflecting a statistically significant increase in neuron production rather than progenitor maintenance (Extended Data Fig. 8c). These findings support the conclusion that loss of CDK5RAP2 leads to premature neural differentiation at the expense of progenitors.

## Discussion

We have established a novel approach to studying human neurodevelopmental processes through *in vitro* culture of cerebral organoids from human pluripotent stem cells. This method recapitulates not only fundamental mechanisms of mammalian neurodevelopment, but also displays characteristics of human brain development. We are hopeful that this method will allow for the study of a variety of neurodevelopmental processes specific to human brain development.

Furthermore, a primary goal in neuroscience is to understand the roots of human neurological disease. We have modelled at least some aspects of microcephaly in these cerebral organoids. The finding that progenitor zones in patient-derived tissues display premature neural differentiation at the expense of early progenitors supports a model in which the founder population of RG progenitors fails to properly expand in patient tissues, thereby leading to an overall smaller brain. This may also explain why mouse models have been unable to recapitulate the severity of the disorder in humans. It is proposed that the mouse founder population of neural progenitors do not undergo expansion to the same extent as in humans before the onset of neurogenesis<sup>7</sup>. Thus, a disruption of the founder population in mice would not lead to as severe an effect as that seen in humans. Overall, our findings suggest that we can use this *in vitro* culture system to model aspects of human neurodevelopment and neurological disease and hopefully provide novel insight into the pathogenesis of these disorders.

## METHODS SUMMARY

For cerebral organoid differentiation, pluripotent stem cells were dissociated from mouse embryonic fibroblasts by dispase treatment followed by trypsinization to

generate single cells. In total, 4,500 cells were plated in each well of an ultra-low binding 96-well plate (Corning) in human ES media with low concentration basic fibroblast growth factor (4 ng ml<sup>-1</sup>) and 50 μM Rho-associated protein kinase (ROCK) inhibitor<sup>49</sup> (Calbiochem). Embryoid bodies were fed every other day for 6 days then transferred to low-adhesion 24-well plates (Corning) in neural induction media containing Dulbecco's modified eagle medium (DMEM)/F12, 1:100 N2 supplement (Invitrogen), Glutamax (Invitrogen), minimum essential media-nonessential amino acids (MEM-NEAA) and 1 μg ml<sup>-1</sup> heparin<sup>50</sup> (Sigma). The neuroepithelial tissues were fed every other day for 5 days. On day 11, tissues were transferred to droplets of Matrigel (BD Biosciences) by pipetting into cold Matrigel on a sheet of Parafilm with small 3 mm dimples. These droplets were allowed to gel at 37 °C, removed from the Parafilm and grown in differentiation media containing a 1:1 mixture of DMEM/F12 and Neurobasal containing 1:200 N2 supplement (Invitrogen), 1:100 B27 supplement without vitamin A (Invitrogen), 3.5 μl l<sup>-1</sup> 2-mercaptoethanol, 1:4,000 insulin (Sigma), 1:100 Glutamax (Invitrogen), 1:200 MEM-NEAA. After 4 days of stationary growth, the droplets were transferred to a spinning bioreactor containing differentiation media as above, except B27 supplement with vitamin A (Invitrogen) was used.

**Online Content** Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 19 December 2012; accepted 2 August 2013.

Published online 28 August 2013.

- Götz, M. & Huttner, W. B. The cell biology of neurogenesis. *Nature Rev. Mol. Cell Biol.* **6**, 777–788 (2005).
- Zecevic, N., Chen, Y. & Filipovic, R. Contributions of cortical subventricular zone to the development of the human cerebral cortex. *J. Comp. Neurol.* **491**, 109–122 (2005).
- Fietz, S. A. *et al.* OSVZ progenitors of human and ferret neocortex are epithelial-like and expand by integrin signaling. *Nature Neurosci.* **13**, 690–699 (2010).
- Hansen, D. V., Lui, J. H., Parker, P. R. L. & Kriegstein, A. R. Neurogenic radial glia in the outer subventricular zone of human neocortex. *Nature* **464**, 554–561 (2010).
- Smart, I. H. M., Dehay, C., Giroud, P., Berland, M. & Kennedy, H. Unique morphological features of the proliferative zones and postmitotic compartments of the neural epithelium giving rise to striate and extrastriate cortex in the monkey. *Cereb. Cortex* **12**, 37–53 (2002).
- Shitamukai, A., Konno, D. & Matsuzaki, F. Oblique radial glial divisions in the developing mouse neocortex induce self-renewing progenitors outside the germinal zone that resemble primate outer subventricular zone progenitors. *J. Neurosci.* **31**, 3683–3695 (2011).
- Lui, J. H., Hansen, D. V. & Kriegstein, A. R. Development and evolution of the human neocortex. *Cell* **146**, 18–36 (2011).
- Fietz, S. A. & Huttner, W. B. Cortical progenitor expansion, self-renewal and neurogenesis — a polarized perspective. *Curr. Opin. Neurobiol.* **21**, 23–35 (2011).
- Cox, J., Jackson, A. P., Bond, J. & Woods, C. G. What primary microcephaly can tell us about brain growth. *Trends Mol. Med.* **12**, 358–366 (2006).
- Megraw, T. L., Sharkey, J. T. & Nowakowski, R. S. Cdk5rap2 exposes the centrosomal root of microcephaly syndromes. *Trends Cell Biol.* **21**, 470–480 (2011).
- Barrera, J. A. *et al.* CDK5RAP2 regulates centriole engagement and cohesion in mice. *Dev. Cell* **18**, 913–926 (2010).
- Lizarraga, S. B. *et al.* Cdk5rap2 regulates centrosome function and chromosome segregation in neuronal progenitors. *Development* **137**, 1907–1917 (2010).
- Pulvers, J. N. *et al.* Mutations in mouse *Aspm* (abnormal spindle-like microcephaly associated) cause not only microcephaly but also major defects in the germline. *Proc. Natl Acad. Sci. USA* **107**, 16595–16600 (2010).
- Gruber, R. *et al.* MCPH1 regulates the neuroprogenitor division mode by coupling the centrosomal cycle with mitotic entry through the Chk1-Cdc25 pathway. *Nature Cell Biol.* **13**, 1325–1334 (2011).
- Sato, T. *et al.* Single Lgr5 stem cells build crypt-villus structures *in vitro* without a mesenchymal niche. *Nature* **459**, 262–265 (2009).
- Suga, H. *et al.* Self-formation of functional adenohipophysis in three-dimensional culture. *Nature* **480**, 57–62 (2011).
- Nakano, T. *et al.* Self-formation of optic cups and storable stratified neural retina from human ESCs. *Cell Stem Cell* **10**, 771–785 (2012).
- Eiraku, M. *et al.* Self-organizing optic-cup morphogenesis in three-dimensional culture. *Nature* **472**, 51–56 (2011).
- Eiraku, M. & Sasai, Y. Self-formation of layered neural structures in three-dimensional culture of ES cells. *Curr. Opin. Neurobiol.* **22**, 768–777 (2012).
- Eiraku, M. *et al.* Self-organized formation of polarized cortical tissues from ESCs and its active manipulation by extrinsic signals. *Cell Stem Cell* **3**, 519–532 (2008).
- Danjo, T. *et al.* Subregional specification of embryonic stem cell-derived ventral telencephalic tissues by timed and combinatory treatment with extrinsic signals. *J. Neurosci.* **31**, 1919–1933 (2011).
- Muguruma, K. *et al.* Ontogeny-recapitulating generation and tissue integration of ES cell-derived Purkinje cells. *Nature Neurosci.* **13**, 1171–1180 (2010).
- Mariani, J. *et al.* Modeling human cortical development *in vitro* using induced pluripotent stem cells. *Proc. Natl Acad. Sci. USA* **109**, 12770–12775 (2012).
- Xia, X. & Zhang, S.-C. Differentiation of neuroepithelia from human embryonic stem cells. *Methods Mol. Biol.* **549**, 51–58 (2009).



25. Swanson, L. W. Mapping the human brain: past, present, and future. *Trends Neurosci.* **18**, 471–474 (1995).
26. Bedogni, F. *et al.* Tbr1 regulates regional and laminar identity of postmitotic neurons in developing neocortex. *Proc. Natl Acad. Sci. USA* **107**, 13129–13134 (2010).
27. Yoo, A. S. *et al.* MicroRNA-mediated conversion of human fibroblasts to neurons. *Nature* **476**, 228–231 (2011).
28. Lessard, J. *et al.* An essential switch in subunit composition of a chromatin remodeling complex during neural development. *Neuron* **55**, 201–215 (2007).
29. Willardsen, M. I. & Link, B. A. Cell biological regulation of division fate in vertebrate neuroepithelial cells. *Dev. Dyn.* **240**, 1865–1879 (2011).
30. Chenn, A. & McConnell, S. K. Cleavage orientation and the asymmetric inheritance of Notch1 immunoreactivity in mammalian neurogenesis. *Cell* **82**, 631–641 (1995).
31. Konno, D. *et al.* Neuroepithelial progenitors undergo LGN-dependent planar divisions to maintain self-renewability during mammalian neurogenesis. *Nature Cell Biol.* **10**, 93–101 (2008).
32. Yingling, J. *et al.* Neuroepithelial stem cell proliferation requires LIS1 for precise spindle orientation and symmetric division. *Cell* **132**, 474–486 (2008).
33. Postiglione, M. P. *et al.* Mouse inscuteable induces apical-basal spindle orientation to facilitate intermediate progenitor generation in the developing neocortex. *Neuron* **72**, 269–284 (2011).
34. Smart, I. H. Proliferative characteristics of the ependymal layer during the early development of the mouse neocortex: a pilot study based on recording the number, location and plane of cleavage of mitotic figures. *J. Anat.* **116**, 67–91 (1973).
35. Zamenhof, S. Quantitative studies of mitoses in fetal rat brain: orientations of the spindles. *Brain Res.* **428**, 143–146 (1987).
36. LaMonica, B. E., Lui, J. H., Hansen, D. V. & Kriegstein, A. R. Mitotic spindle orientation predicts outer radial glial cell generation in human neocortex. *Nature Commun.* **4**, 1665 (2013).
37. Hevner, R. F. *et al.* Tbr1 regulates differentiation of the preplate and layer 6. *Neuron* **29**, 353–366 (2001).
38. Shafit-Zagardo, B. & Kalcheva, N. Making sense of the multiple MAP-2 transcripts and their role in the neuron. *Mol. Neurobiol.* **16**, 149–162 (1998).
39. Frotscher, M. Cajal-Retzius cells, Reelin, and the formation of layers. *Curr. Opin. Neurobiol.* **8**, 570–575 (1998).
40. Tsai, L.-H. & Gleeson, J. G. Nucleokinesis in neuronal migration. *Neuron* **46**, 383–388 (2005).
41. Gaspard, N. *et al.* An intrinsic mechanism of corticogenesis from embryonic stem cells. *Nature* **455**, 351–357 (2008).
42. De Carlos, J. A. & O'Leary, D. D. Growth and targeting of subplate axons and establishment of major cortical pathways. *J. Neurosci.* **12**, 1194–1211 (1992).
43. Chédotal, A. Further tales of the midline. *Curr. Opin. Neurobiol.* **21**, 68–75 (2011).
44. Sato, T. R., Gray, N. W., Mainen, Z. F. & Svoboda, K. The functional microarchitecture of the mouse barrel cortex. *PLoS Biol.* **5**, e189 (2007).
45. Bond, J. *et al.* A centrosomal mechanism involving CDK5RAP2 and CENPJ controls brain size. *Nature Genet.* **37**, 353–355 (2005).
46. Pagnamenta, A. T. *et al.* A novel nonsense *CDK5RAP2* mutation in a Somali child with primary microcephaly and sensorineural hearing loss. *Am. J. Med. Genet.* **158A**, 2577–2582 (2012).
47. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006).
48. Okita, K., Ichisaka, T. & Yamanaka, S. Generation of germline-competent induced pluripotent stem cells. *Nature* **448**, 313–317 (2007).
49. Watanabe, K. *et al.* A ROCK inhibitor permits survival of dissociated human embryonic stem cells. *Nature Biotechnol.* **25**, 681–686 (2007).
50. Hu, B.-Y. & Zhang, S.-C. Directed differentiation of neural-stem cells and subtype-specific neurons from hESCs. *Methods Mol. Biol.* **636**, 123–137 (2010).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We are grateful to members of the Knoblich laboratory for technical expertise and feedback, A. Peer, P. Moeseneder and N. Corsini for experimental support and M. Repic for help with establishing organoid electroporations. We also thank the Stem Cell and BioOptics core facilities of IMBA/IMP for technical support. We would especially like to thank the patients and their families for participating in this study. We would also like to thank S. McGurk for providing control MRI images. M.A.L. received funding from an EMBO post-doctoral fellowship and a Helen Hay Whitney post-doctoral fellowship. Work in A.P.J.'s laboratory is supported by the Medical Research Council, a starter grant from the European Research Council (ERC) and the Lister Institute for Preventative Medicine. This research was also supported in part by Wellcome Trust grant WT098051. Work in J.A.K.'s laboratory is supported by the Austrian Academy of Sciences, the Austrian Science Fund (FWF) (projects Z153-B09 and I552-B19) and an advanced grant from ERC.

**Author Contributions** M.A.L. and J.A.K. conceived the project and experimental design and wrote the manuscript. M.A.L. performed experiments and analysed data. M.R., C.-A.M. and D.W. performed experiments and analysed data under the supervision of J.A.K., J.M.P. and A.P.J. L.S.B., M.E.H. and T.H. performed patient diagnosis and provided MRIs coordinated by A.P.J. J.A.K. directed and supervised the project.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.A.K. ([juergen.knoblich@imba.oeaw.ac.at](mailto:juergen.knoblich@imba.oeaw.ac.at)).



## METHODS

**Plasmid constructs and materials.** GFP plasmid used for co-electroporation with shRNA and for live imaging was pCAG-GFP (Addgene plasmid 11150)<sup>51</sup>. shRNAs targeting human *CDK5RAP2* were cloned using pSuper shRNA expression strategy (OligoEngine). Targeting sequences were as follows: shRNA 1, 5'-AGGA CGTGTGCTTCAGAAAT-3'; shRNA 2, 5'-AGAGTCAGCCTTCTGCTAAAG-3'; shRNA 3, 5'-GTGGAAGATCTCCTAACTAAA-3'; shRNA 4, 5'-ACTATGAGA CTGCTCTATCAG-3'. The *CDK5RAP2* expression construct was generated using the Gateway system (Invitrogen) by PCR amplification of *CDK5RAP2* from MGC human *CDK5RAP2* cDNA (clone ID: 9052276) using the primers with AttB sites: forward, 5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGATGGACT TGGTGTGGAAGA-3'; reverse, 5'-GGGGACCACTTTGTACAAGAAAGCTG GGTACAGTTTATTGGCTGAAAGTTCTTCTC-3'. *CDK5RAP2* was cloned into destination vector pcDNA3.1/nV5-DEST.

**Cerebral organoid culture conditions.** Human H9 ES (WA09) cells were obtained from WiCell at passage 26 with verified normal karyotype and contamination-free. iPS cells were obtained from System Biosciences (SC101A-1) verified pluripotent and contamination-free. All human pluripotent stem cell lines were regularly checked and confirmed negative for mycoplasma. Pluripotent stem cells were maintained on CF-1-gamma-irradiated mouse embryonic stem cells (MEFs) (Global Stem) according to WiCell protocols. On day 0 of organoid culture, ES cells or iPS cells at less than passage 50 were dissociated from MEFs by dispase treatment and MEFs were removed by gravity separation of stem cell colonies from MEFs before trypsinization of stem cells to generate single cells. In total, 4,500 cells were then plated in each well of an ultra-low-binding 96-well plate (Corning) in human ES media with low concentration basic fibroblast growth factor (4 ng ml<sup>-1</sup>) and 50 μM Rho-associated protein kinase (ROCK) inhibitor<sup>49</sup> (Calbiochem).

Embryoid bodies were fed every other day for 6 days then transferred to low-adhesion 24-well plates (Corning) in neural induction media containing Dulbecco's modified eagle medium (DMEM)/F12, 1:100 N2 supplement (Invitrogen), Glutamax (Invitrogen), minimum essential media-nonessential amino acids (MEM-NEAA) and 1 μg ml<sup>-1</sup> heparin<sup>50</sup> (Sigma). These began forming neuroepithelial tissues, which were fed every other day for 5 days. On day 11 of the protocol, tissues were transferred to droplets of Matrigel (BD Biosciences) by pipetting into cold Matrigel on a sheet of Parafilm with small 3 mm dimples. These droplets were allowed to gel at 37 °C and were subsequently removed from the Parafilm and grown in differentiation media containing a 1:1 mixture of DMEM/F12 and Neurobasal containing 1:200 N2 supplement (Invitrogen), 1:100 B27 supplement without vitamin A (Invitrogen), 3.5 μl l<sup>-1</sup> 2-mercaptoethanol, 1:4,000 insulin (Sigma), 1:100 Glutamax (Invitrogen) and 1:200 MEM-NEAA.

After 4 days of stationary growth, the tissue droplets were transferred to a spinning bioreactor containing differentiation media as above except B27 supplement with vitamin A (Invitrogen) was used. Because retinoic acid has been shown to be important for neuronal differentiation *in vivo*<sup>52</sup>, we included it in the final media used to differentiate the cerebral organoids.

**Mouse organoid culture conditions.** Mouse A9 ES cells were cultured on mitomycin-C-growth-inactivated MEFs and passaged according to standard protocols<sup>53</sup>. For the generation of mouse organoids, the organoid protocol was applied with the following modifications: cells were trypsinized and 2,000 stem cells were plated in each well of an ultra-low-binding 96-well plate in differentiation medium as described in ref. 20 (medium containing 10 μM SB431542 but without DKK1). Subsequent steps were followed according to the human organoid method using identical media compositions, with the exception that for mouse tissues faster timing was used according to morphology. Embryoid bodies were transferred to neural induction medium on day 4, embedded in Matrigel droplets on day 6, and on day 9 transferred to the spinning bioreactor.

**Organoid electroporation.** Electroporation was performed using a Petri dish tissue electrode and electro-square-porator (ECM 830), both from BTX Harvard Apparatus. A total of 3 μl of 2 μg μl<sup>-1</sup> total plasmid (GFP for live imaging, 1.8 μg μl<sup>-1</sup> shRNA plus 0.2 μg μl<sup>-1</sup> GFP for shRNA experiments) was injected in 4–5 locations within the organoid and electroporation was performed in differentiation media without antibiotics at 5 pulses, 80 V, 50-ms duration, 1-s interval. For rescue experiments, GFP expression plasmid and the *CDK5RAP2* construct were co-electroporated at equal concentrations (1 μg μl<sup>-1</sup> each).

**Live imaging in organoids.** Live imaging was performed using a LSM780 confocal laser scanning system (Zeiss) equipped with temperature and CO<sub>2</sub> control. For calcium imaging, Fluo-4 direct (Life Technologies) was prepared according to manufacturer and applied 60 min before the start of imaging. Imaging was performed at 494 nm excitation and 516 nm emission, with frames taken every 20 s for 100 frames. Data analysis of calcium imaging was performed using ImageJ (Fiji). ROIs were manually selected and mean fluorescence was calculated for each time frame. Change in fluorescence was calculated as follows:  $\Delta F/F = (F - F_{\text{basal}})/F_{\text{background}}$ , in which  $F_{\text{background}}$  was the lowest mean fluorescence value across imaging

and  $F_{\text{background}}$  was the average mean fluorescence across all frames. Glutamate was added by bath application to media during imaging at a final concentration 100 μM. TTX was added by bath application to media during imaging at a final concentration of 1 μM and imaging was resumed after a 10-min incubation time.

**Histology and immunofluorescence.** Tissues were fixed in 4% paraformaldehyde for 20 min at 4 °C followed by washing in PBS three times for 10 min. Tissues were allowed to sink in 30% sucrose overnight and then embedded in 10%/7.5% gelatin/sucrose and cryosectioned at 20 μm. Tissue sections were stained with haematoxylin and eosin or used for immunostaining. For immunohistochemistry, sections were blocked and permeabilized in 0.25% Triton X-100 and 4% normal donkey serum in PBS. Sections were then incubated with primary antibodies in 0.1% Triton X-100, 4% normal donkey serum at the following dilutions: N-cadherin (mouse, BD Biosciences 610920, 1:500), SOX2 (rabbit, Chemicon, AB5603, 1:300), TUJ1 (mouse, Covance MMS-435P, 1:750), TUNEL (*In situ* Cell Death Detection Kit-Fluorescein, Roche), FOXG1 (rabbit, Abcam ab18259, 1:200), EMX1 (rabbit, Sigma HPA006421, 1:50), KROX20 (rabbit, Covance PRB-236P, 1:100), PAX2 (mouse, Abnova H00005076-M01, 1:200), LMO4 (goat, Santa Cruz sc-11122, 1:50), TSHZ2 (rabbit, Sigma SAB4500379, 1:50), OTX1+2 (rabbit, Abcam ab21990, 1:200), GBX2 (goat, Santa Cruz sc22230, 1:100), AUTS2 (rabbit, Sigma HPA000390, 1:250), NKX2-1 (rabbit, Epitomics 6594-1, 1:250), PAX6 (mouse monoclonal, DSHB, 1:200), PAX6 (rabbit, Covance PRB-278P, 1:300), calretinin (mouse, Swant 6B3, 1:100), NRP2 (goat, RandD systems AF2215, 1:40), FZD9 (rabbit, Acris SP4153P, 1:200), PROX1 (mouse, Chemicon MAB5654, 1:200), TTR (sheep, AbD Serotec AHP1837, 1:100), TBR2 (rabbit, Chemicon AB9618, 1:500), TBR1 (rabbit, Abcam ab31940, 1:300), MAP2 (mouse, Chemicon MAB3418, 1:300), H3S10ph (rabbit, Cell Signaling Technology 9706S, 1:300), p-vimentin (mouse, MBL International D076-3S, 1:250), BrdU (rat, AbD Serotec OBT0030CX, 1:500), BAF53A (rabbit, Bethyl IHC-00287, 1:250), BAF53B (rabbit, Abcam ab140642, 1:250), reelin (mouse, Millipore MAB5366, 1:200), CTIP2 (rat, Abcam ab18465, 1:100), SATB2 (rabbit, Abcam ab34735, 1:100), DCX (goat, Santa Cruz sc-8066, 1:300), BRN2 (goat, Santa Cruz sc-6029, 1:40). Secondary antibodies used were donkey Alexa Fluor 488, 568 and 647 conjugates (Invitrogen, 1:500). For sections stained for BrdU, sections were first incubated with 2N HCl at 37 °C for 20 min followed by washing three times in PBS before blocking.

**RT-PCR.** Total messenger RNA samples were isolated from whole organoids or human ES cells in triplicate using Trizol reagent (Invitrogen). Potential contaminating DNA was removed using DNA-Free (Ambion) and 1 μg RNA was used for cDNA synthesis using SuperScript III (Life Technologies). PCR conditions and number of cycles (25–35 cycles) for each primer pair were empirically determined using human ES cDNA or human fetal brain cDNA (Invitrogen). Cycles were run at 94 °C denaturation for 30 s, 58–62 °C annealing for 45 s, depending on primer pair, and 72 °C extension for 30 s. Primer pairs used were as follows: OCT4A forward: 5'-GGAGAAGCTGGAGCAAAACC-3', reverse: 5'-TGGCTGAATACCTTCCC AAA-3'; NANOG forward: 5'-GATTGTGGGCGCTGAAGAAA-3', reverse: 5'-CTTTGGGACTGGTGAAGAA-3'; SOX1 forward: 5'-TATCTTCTGCTCCGG CTGTT-3', reverse: 5'-GGGTCTTCCCTTCTCTCCTC-3'; PAX6 forward: 5'-AG TTCTTCGCAACCTGGCTA-3', reverse: 5'-ATTCTCTCCCTCTCTCTCT-3'; ACTB forward: 5'-AAATCTGCACCACACCTTC-3', reverse: 5'-AGAGGCGT ACAGGGATAGCA-3'; FOXG1 forward: 5'-AGGAGGGCGAGAAGAAGAAC-3', reverse: 5'-TGAACCTGTAGATGCCGTG-3'; SIX3 forward: 5'-CTATCAACA ACCCCCAACCA-3', reverse: 5'-AGCCGTGCTGTCTCTAGAAA-3'; KROX20 forward: 5'-TTGACCAGATGAACGGAGTG-3', reverse: 5'-CTTGCCCATGTA AGTGAAGGT-3'; ISL1 forward: 5'-GCTTTGTTAGGGATGGGAAA-3', reverse: 5'-ACTCGATGTGATACACCTTGA-3'.

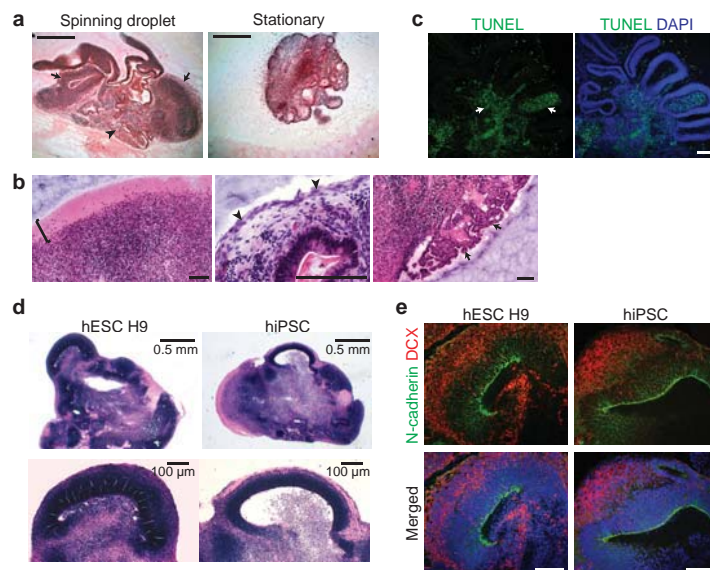
**Cell culture and western blot.** HEK293T cells were grown in 10% FBS/DMEM and split at 40% into a 6-well dish (BD Falcon) followed by transfection the next day using TurboFect (Thermo Scientific) with 5 μg plasmid DNA. Cells were lysed 2 days later and western blot was performed using rabbit anti-CDK5RAP2 (A300-554A, Bethyl Laboratories, 1:10,000) followed by blotting for mouse anti-α-tubulin (mouse, Sigma T6199, 1:10,000). Dermal fibroblasts were obtained by skin punch biopsy and were cultured in amnioMAX C-100 complete medium (Invitrogen) and maintained in a 37 °C incubator with 5% CO<sub>2</sub> and 3% O<sub>2</sub>. Cells were lysed in 50 mM Tris-HCl, pH 8, 280 mM NaCl, 0.5% NP40, 0.2 mM EDTA, 0.2 mM EGTA and 10% glycerol supplemented with protease inhibitor tablet (Roche). Protein samples were run on a 3–8% Tris-acetate gel (Invitrogen) followed by immunoblotting using rabbit anti-CDK5RAP2 (A300-554A, Bethyl Laboratories, 1:2,000) and mouse anti-vinculin (V9264, Sigma, 1:2,000). To perform immunofluorescence, patient fibroblasts were fixed in -20 °C methanol for 7 min and then blocked in PBS/1% bovine serum albumin. Cells were then incubated in rabbit anti-CDK5RAP2 (A300-554A, Bethyl Laboratories, 1:2,000) and mouse anti-CPAP (SC-81432, Santa Cruz Biotechnology, 1:100) in blocking solution. Secondary antibodies used were donkey Alexa Fluor 488 and 568 conjugates (Invitrogen, 1:500).

**Research subject and gene identification.** Genomic DNA was extracted from peripheral blood of patient 3842 and the patient's parents by standard methods. Informed consent was obtained from the family and the study approved by the Multi-centre Research Ethics Committee for Scotland (04:MRE00/19). Whole-exome capture and sequencing was performed at the Wellcome Trust Sanger Institute, UK. DNA was sheared to 150-base-pair lengths by sonification (Covaris) before whole-exome capture and amplification using the SureSelect Human All Exon 50Mb kit (Agilent). Fragments were sequenced using the Illumina HiSeq platform. 76-base-pair paired-end sequence reads were aligned to the UCSC genome browser hg19 reference sequence using BWA. Sequence variants were obtained using GenomeAnalysisTK (<http://www.broadinstitute.org/gatk/>) and annotated with transcript and protein consequence, polyphen, condel and SIFT scores. Mutations were confirmed by bi-directional sequencing of PCR products using dye terminator chemistry on an ABI 3730 capillary sequencer (Applied Biosystems).

**Patient iPS cell reprogramming.** Patient skin fibroblasts were reprogrammed using lentiviral delivery of *OCT4*, *SOX2*, *KLF4* and *MYC*. Lentivirus production: a DNA mix consisting of virus packaging vectors (tat, rev, gag/pol, 1.5 µg each, and vsv-g, 3 µg) and the loxP-flanked OKSM reprogramming vector (*OCT4*, *KLF4*, *SOX2*, *MYC*, 30 µg) were transfected into 293 cells. In brief, 112.5 µl Fugene6 was added dropwise to 2 ml DMEM under constant vortexing followed by a 10-min incubation at room temperature (22 °C). The DNA mix was added to the DMEM/Fugene6 mix while vortexing to generate the final transfection mix. After a 15-min incubation at room temperature, the transfection mix was added onto 80% confluent 293 cells, cultured in 13 ml 293 culture medium. Virus-containing medium

was collected and replaced with fresh medium 48, 60 and 72 h after transfection. The viral supernatant was stored at 4 °C. Reprogramming of human dermal fibroblasts:  $1 \times 10^5$  dermal fibroblasts were seeded the day before infection onto 10-cm and 6-cm 0.1% gelatin-coated culture dishes. Cells were incubated for 12 h with viral supernatant 1:1 mixed with dermal fibroblast medium supplemented with  $4 \mu\text{g ml}^{-1}$  polybrene. Thereafter, cells were washed with  $1 \times$  PBS and cultured for 2 more days in dermal fibroblast medium. After 2 days medium was switched to human iPS cell medium supplemented with  $10 \text{ ng ml}^{-1}$  basic fibroblast growth factor (peprotech, cat. no. 100-18B),  $10 \mu\text{M}$  CHIR99021 (stemgent, cat. no. 04-0004) and  $1 \mu\text{M}$  PD 0325901 (stemgent, cat. no. 04-0006) and cells cultured for 21 days. Medium was changed every day. Outgrowing colonies, identified by morphological appearance, were picked and passaged on inactivated CF-1 MEFs (global stem, cat. no. GSC-6201M). Patient-derived iPS lines were compared to control iPS cells obtained from a healthy donor (System Biosciences, SC101A-1). Alkaline phosphatase staining was performed using Vector Blue Alkaline Phosphatase Substrate Kit (Vector Laboratories, SK5300). Quantifications in patient and control iPS-cell-derived organoids were performed blinded using coded file names in ImageJ.

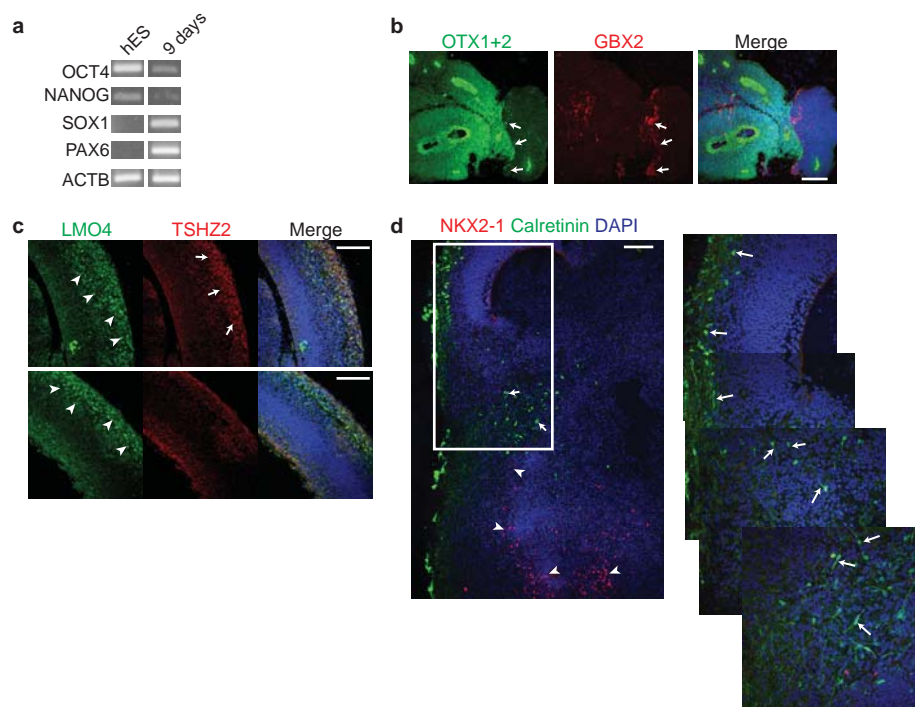
51. Matsuda, T. & Cepko, C. L. Electroporation and RNA interference in the rodent retina *in vivo* and *in vitro*. *Proc. Natl Acad. Sci. USA* **101**, 16–22 (2004).
52. Siegenthaler, J. A. *et al.* Retinoic acid from the meninges regulates cortical neuron generation. *Cell* **139**, 597–609 (2009).
53. Tremml, G., Singer, M. & Malavarca, R. Culture of mouse embryonic stem cells. *Curr. Protoc. Stem Cell Biol.* **1**, Unit-1C.4 (2008).



**Extended Data Figure 1 | Generation of cerebral organoids from multiple human pluripotent stem cells.** **a**, Haematoxylin and eosin staining of cerebral organoids compared with stationary culture reveals overall larger tissues with substructure reminiscent of brain regions such as forebrain cortex (arrows) and choroid plexus (arrowhead). **b**, Higher magnification images of haematoxylin and eosin stained organoids revealing layering reminiscent of the cerebral cortical molecular layer (bar), as well as tissue reminiscent of meninges (arrowheads) and choroid plexus (arrows). **c**, TUNEL staining (green) revealing cell death in the interior regions (arrows) of the cerebral organoid

with cortical regions developing along the exterior. DAPI marks nuclei (blue). **d**, Haematoxylin and eosin staining of organoids generated from human H9 ES cells as well as human iPS cells display similar size and complex morphology as well as the presence of advanced forebrain tissues, shown at higher magnification in the bottom panels. **e**, Staining for N-cadherin (green) and newborn neurons (DCX, red) in tissues generated from both human H9 ES cells and human iPS cells reveals similar organization and intact apical basal polarity in both types of tissues. Scale bars, 0.5 mm (**a**), 100  $\mu$ m (**b**, **c**, **d** bottom panels, **e**), and 0.5 mm (**d** top panels).

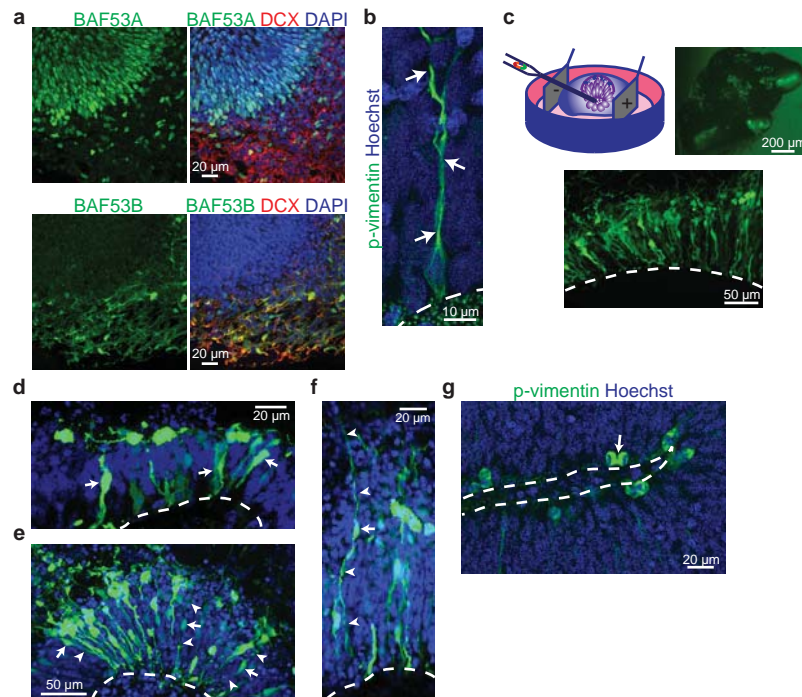




**Extended Data Figure 2 | Neural identity during differentiation of cerebral organoids.** **a**, RT-PCR for the pluripotency markers OCT4 and NANOG as well as neural identity markers SOX1 and PAX6 in undifferentiated human ES cells and following differentiation at 9 days, revealing induction of neural identity with decreased pluripotent identity at 9 days of differentiation.

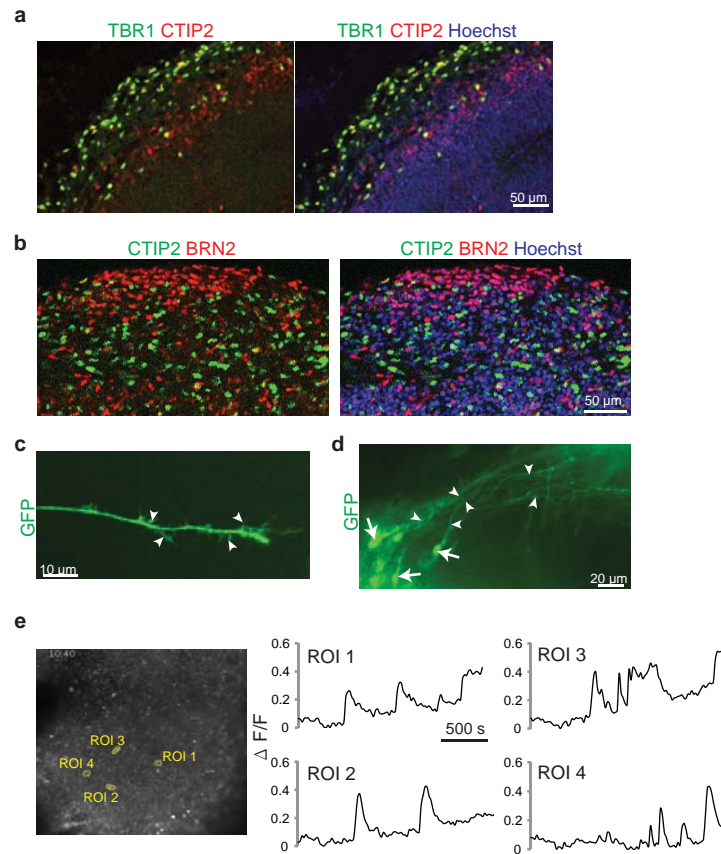
**b**, Immunohistochemistry for the forebrain/midbrain marker OTX1/2 (green) and the hindbrain marker GBX2 (red) at 16 days of differentiation, revealing primarily fore/midbrain identity with adjacent regions of hindbrain reminiscent of the mid-hindbrain boundary (arrows). DAPI marks nuclei (blue). **c**, Staining for the cortical lobe markers LMO4 (frontal and occipital marker, green) and TSHZ2 (occipital marker, red). Note the expected nuclear

staining (arrows and arrowheads) for both in one region (top panels) suggesting occipital identity, whereas only LMO4 staining (arrowheads) is clearly evident in another region (bottom panels), suggesting frontal identity. DAPI marks nuclei (blue). **d**, Staining for the ventral marker NKX2-1 (red) and the cortical interneuron marker calretinin (green) on an organoid containing both ventral (arrowheads) and dorsal (top left) regions within one section. Images on the right are higher magnification stitched images of the region outlined in the lower magnification image at left. Calretinin interneurons can be seen between the two regions with typical morphology of migration and redirection towards the dorsal cortex (arrows). Scale bars, 100  $\mu$ m.



**Extended Data Figure 3 | RG organization and morphology.** **a**, Staining for the chromatin remodelling BAF components BAF53A (also known as ACTL6A) (green, top panels) and BAF53B (also known as ACTL6B) (green, bottom panels) in serial sections of the same tissue showing the neural-progenitor-specific BAF53A expressed in VZ RGs, whereas the neuron-specific BAF53B is expressed in DCX<sup>+</sup> (red) neurons outside the VZ. **b**, Higher magnification image of p-vimentin staining (green) of a dividing radial glia revealing the long basal process typical of radial glial morphology. **c**, Schematic of electroporation technique. Plasmid DNA was injected into fluid-filled cavities within the organoid and an electric pulse was applied to electroporate

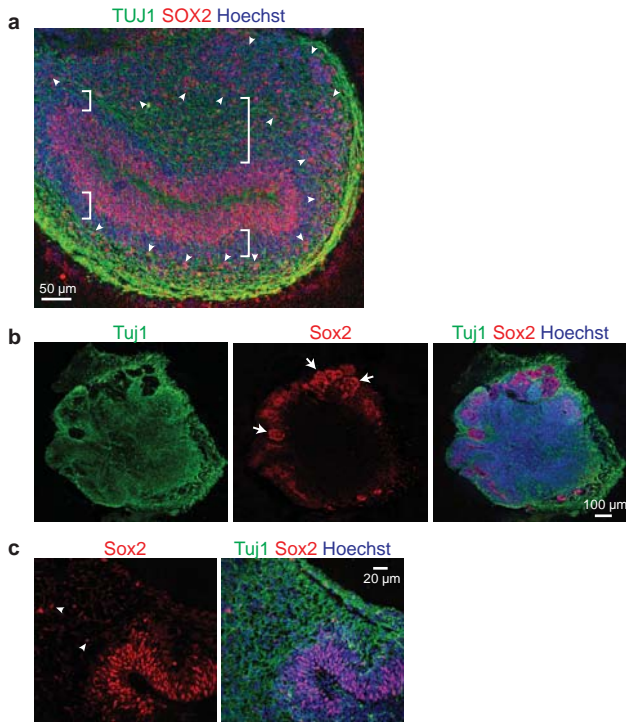
cells (radial glial progenitors) adjacent to the cavity. This resulted in several regions of electroporation (top right, GFP in green) and high efficiency of electroporation of RGs (bottom, GFP in green). **d**, GFP-electroporated progenitors (arrows) in an early stage tissue (18 days) revealing neuroepithelial morphology. **e**, GFP-electroporated tissue at 30 days, revealing RG (arrows) with typical bipolar morphology (arrowheads). **f**, GFP-electroporated tissue at 36 days revealing more advanced thicker cortical region with RG (arrow) exhibiting long apical and basal processes (arrowheads). **g**, p-Vimentin (green) staining revealing a mitotic cell at the apical surface during anaphase (arrow) with a horizontal orientation of division.



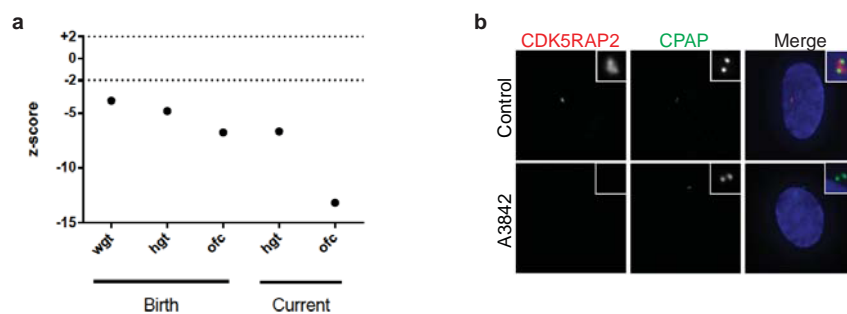
**Extended Data Figure 4 | Spatial organization and characteristics of cortical neuron identities.** **a**, Staining for the preplate marker TBR1 (green) and the deep-layer marker CTIP2 (red) at day 30, revealing rudimentary spatial separation reminiscent of the early stages of cortical plate development. **b**, Immunohistochemistry for the early-born neuron marker CTIP2 (green) and later-born neuron marker BRN2 (red) reveals independent neuron populations exhibiting rudimentary separation at 30 days of differentiation.

**c**, GFP (green)-electroporated neuronal axon 5 days after electroporation displaying complex morphology and axon branching (arrowheads). **d**, GFP (green)-electroporated cortical neurons (arrows) 5 days after electroporation extend long-range axons with evidence of axon bundling (arrowheads) similar to that seen in pyramidal tracts. **e**, Single-cell tracings of calcium surges in individual neurons (ROI, outlined in left panel) as measured by change in fluorescence (arbitrary units).



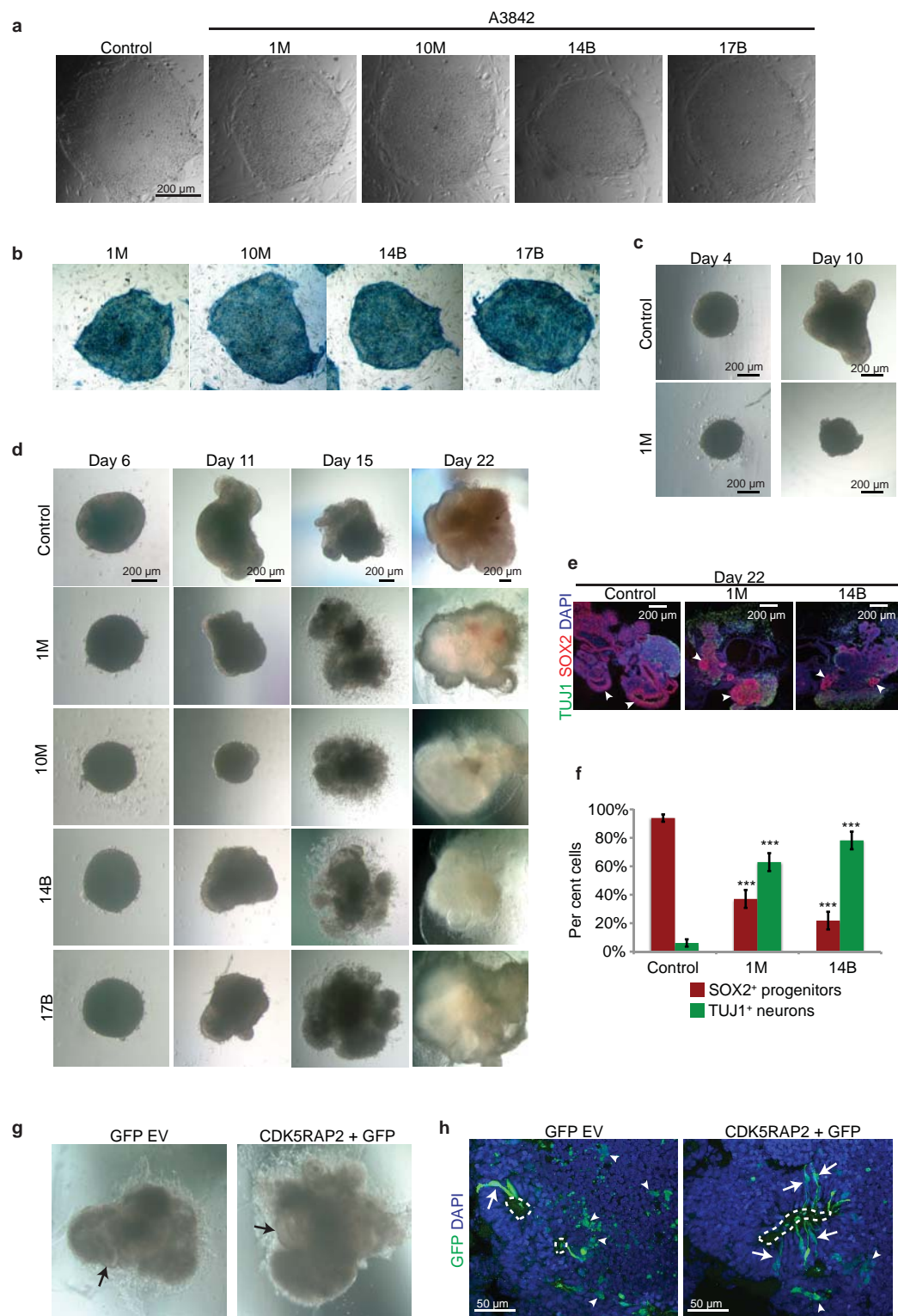


**Extended Data Figure 5 | Human features of cortical development are not recapitulated in mouse organoids.** **a**, Low-magnification image of the region shown in Fig. 5a revealing the presence of a separated region of oRGs (demarcated by arrowheads) that appear separate from the VZ in all regions (brackets) but more separated and with a layer of TUJ1<sup>+</sup> fibres in between in thicker parts of the cortical tissue (larger bracket). The entire organoid can be seen in Fig. 1c. **b**, Low-magnification image of a cerebral organoid derived from mouse ES cells stained for neurons (Tuj1, green) and neural progenitors (Sox2, red) revealing overall smaller organoid size as well as smaller cortical regions (arrows) than in humans. **c**, Higher magnification of a region of cortical identity in mouse cerebral organoids stained for RG progenitors (Sox2, red) revealing the presence of only a few oRGs (arrowheads) that do not organize into a separate layer such as that seen in humans.



**Extended Data Figure 6 | Patient growth parameters.** **a**, All growth parameters were significantly reduced both at birth and postnatally, with all z-scores less than  $-2$  s.d. from the population mean for age and sex (dashed line). Weight (wgt), height (hgt) and head circumference (occipitofrontal circumference, ofc) at birth and at current age of 3.5 years of age. Head

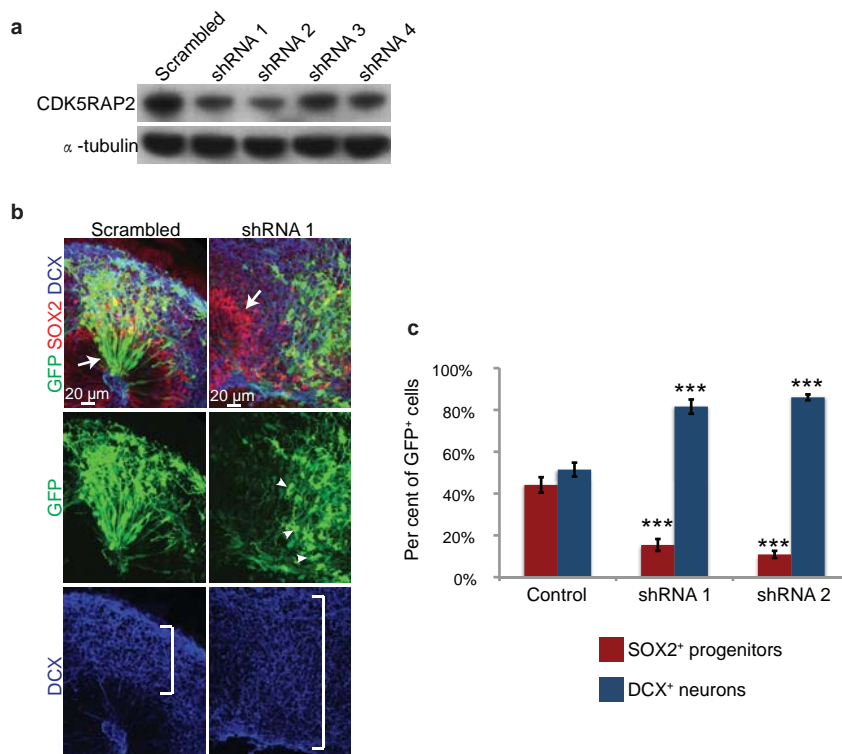
circumference was much more severely affected than height and weight, indicating that brain volume was disproportionately reduced as a result of more severe growth restriction. **b**, CDK5RAP2 (red) is absent from the centrosome in patient fibroblasts. Immunofluorescent microscopy images of patient (A3842) and control cells, also stained with the centriolar marker CPAP (green).





**Extended Data Figure 7 | Characterization of patient-derived iPS cells and cerebral organoids.** **a**, iPS cells derived from A3842 patient skin fibroblasts exhibit typical ES-cell-like morphology. Four lines were chosen for analysis on the basis of this typical morphology and pluripotency. **b**, Alkaline phosphatase staining (blue) of patient-derived iPS cell colonies revealing pluripotency. **c**, Representative early organoid culture of patient (line 1M) and control using the protocol and timing established for normal human ES cells. Patient organoids were much smaller and failed to thrive, therefore the protocol was slightly modified with increased starting cell number to produce neural tissues. **d**, Patient-derived tissues using increased starting cell number displayed neuroepithelium but did not form thick fluid-filled cortical tissues compared with control-derived tissues. Patient-derived tissues also display outgrowth with neural morphology compared with control. **e**, Staining of patient and control organoids at an early stage (day 22) for neurons (TUJ1, green) and RG

(SOX2, red) revealed smaller progenitor zones (arrowheads) and increased neurons in patient-derived tissues (lines 1M and 14B are shown here). **f**, Quantification of the percentage of SOX2<sup>+</sup> progenitors and TUJ1<sup>+</sup> neurons in cerebral cortical regions of control and two lines of patient-derived tissues (1M and 14B) at the early stage of day 22. Error bars are  $\pm$  s.e.m. \*\*\* $P < 0.001$  compared with control, Student's *t*-test.  $n = 4$  tissues for each line. **g**, Bright-field image of patient-derived tissues (line 14B) electroporated with either GFP alone (left panel) or a GFP and CDK5RAP2 expression construct (right panel). Note the presence of larger neuroepithelial tissue (arrows) in CDK5RAP2-electroporated tissue compared with control. **h**, GFP staining (green) in GFP control (left) and CDK5RAP2 co-electroporated patient-derived tissues (14B) revealing the presence of multiple GFP<sup>+</sup> neurons (arrowheads) in control 6 days after electroporation, whereas CDK5RAP2-electroporated tissues display multiple GFP<sup>+</sup> RG (arrows).



**Extended Data Figure 8 | shRNA-mediated knockdown of *CDK5RAP2* in human organoids.** **a**, Western blot for endogenous *CDK5RAP2* in 293T cells transfected with four different shRNAs against *CDK5RAP2*. shRNAs 1 and 2 are most efficient, whereas shRNA 4 leads to a modest reduction in protein.  $\alpha$ -Tubulin is shown as a loading control. **b**, Higher magnification of results in Fig. 6i showing neuronal morphology of *CDK5RAP2* shRNA (GFP, green, arrowheads)-electroporated cells. These exhibit increased DCX (blue)

expression with an expansion of the zone of DCX positivity (bars) and a loss of SOX2 (red) compared with scrambled electroporated or adjacent non-electroporated tissue (arrows). **c**, Quantification of the percentage of GFP<sup>+</sup> electroporated cells exhibiting SOX2<sup>+</sup> progenitor identity or DCX<sup>+</sup> neuronal identity in scrambled control- or shRNA-co-electroporated tissues. \*\*\* $P < 0.001$  compared to control, Student's *t*-test,  $n = 4$  tissues for each shRNA. Error bars are  $\pm$  s.e.m.

# Usp16 contributes to somatic stem-cell defects in Down's syndrome

Maddalena Adorno<sup>1</sup>, Shaheen Sikandar<sup>1</sup>, Siddhartha S. Mitra<sup>1</sup>, Angera Kuo<sup>1</sup>, Benedetta Nicolis Di Robilant<sup>1,2</sup>, Veronica Haro-Acosta<sup>1</sup>, Youcef Ouadah<sup>1,3</sup>, Marco Quarta<sup>4</sup>, Jacqueline Rodriguez<sup>5</sup>, Dalong Qian<sup>1</sup>, Vadiyala M. Reddy<sup>6</sup>, Samuel Cheshier<sup>1,7</sup>, Craig C. Garner<sup>5</sup> & Michael F. Clarke<sup>1</sup>

Down's syndrome results from full or partial trisomy of chromosome 21. However, the consequences of the underlying gene-dosage imbalance on adult tissues remain poorly understood. Here we show that in Ts65Dn mice, which are trisomic for 132 genes homologous to genes on human chromosome 21, triplication of *Usp16* reduces the self-renewal of haematopoietic stem cells and the expansion of mammary epithelial cells, neural progenitors and fibroblasts. In addition, *Usp16* is associated with decreased ubiquitination of Cdkn2a and accelerated senescence in Ts65Dn fibroblasts. *Usp16* can remove ubiquitin from histone H2A on lysine 119, a critical mark for the maintenance of multiple somatic tissues. Downregulation of *Usp16*, either by mutation of a single normal *Usp16* allele or by short interfering RNAs, largely rescues all of these defects. Furthermore, in human tissues overexpression of USP16 reduces the expansion of normal fibroblasts and postnatal neural progenitors, whereas downregulation of USP16 partially rescues the proliferation defects of Down's syndrome fibroblasts. Taken together, these results suggest that USP16 has an important role in antagonizing the self-renewal and/or senescence pathways in Down's syndrome and could serve as an attractive target to ameliorate some of the associated pathologies.

Down's syndrome is a complex clinical syndrome associated with multiple pathological conditions, including deficits in motor skills and learning<sup>1,2</sup>. It has also been associated with early onset and higher incidence of ageing-related phenomena such as Alzheimer's disease<sup>3,4</sup>. Recently, it has been proposed that the ageing process is correlated with an impaired or exhausted ability of stem cells to self-renew<sup>5</sup>, raising the possibility that stem-cell defects have a role in Down's syndrome.

To gain insights into the molecular mechanisms underlying abnormalities of tissue homeostasis in Down's syndrome, we investigated haematopoietic stem cells in two Down's syndrome mouse models, Ts65Dn and Ts1Cje (Fig. 1a). Ts65Dn mice have a third copy of approximately two-thirds of the murine genes homologous to genes on human chromosome 21 and recapitulate several traits of the human disease, including learning and memory deficits<sup>6</sup>. Ts1Cje mice present only 79 of the 132 trisomic genes of Ts65Dn mice (Fig. 1a)<sup>1</sup> and certain defects, such as neurologic impairment, are less severe in these mice<sup>7,8</sup>. One of the genes uniquely triplicated in Ts65Dn is *Usp16*, a deubiquitination enzyme involved in chromatin remodelling and cell cycle progression<sup>9</sup>. Our studies implicated *Usp16* as a major factor causing expansion defects and senescence in multiple somatic tissues of Ts65Dn mice.

## Ts65Dn, but not Ts1Cje, mice have defective HSCs

Consistent with previous observations<sup>10</sup>, transplantation of  $5 \times 10^5$  CD45.2<sup>+</sup> Ts65Dn cells into CD45.1<sup>+</sup> C57BL/6 lethally irradiated mice resulted in lower haematopoietic chimaerism in recipient mice compared to transplantation of control bone marrow cells (Extended Data Fig. 1a). To determine the trisomic genes responsible for the impaired engraftment potential of the Ts65Dn bone marrow cells, we investigated the frequency and properties of haematopoietic stem cells (HSCs) in Ts65Dn and Ts1Cje mice. Immunophenotypical analyses showed that the fraction

of CD150<sup>+</sup> CD48<sup>-</sup> or CD34 Flt3<sup>-</sup> KLS (c-Kit<sup>+</sup> Lineage<sup>-</sup> Sca<sup>+</sup>) cells, both of which enrich for HSCs in mice<sup>11,12</sup>, was reduced by more than threefold in Ts65Dn mice, but not in Ts1Cje mice (Fig. 1b, c and Extended Data Fig. 1b).

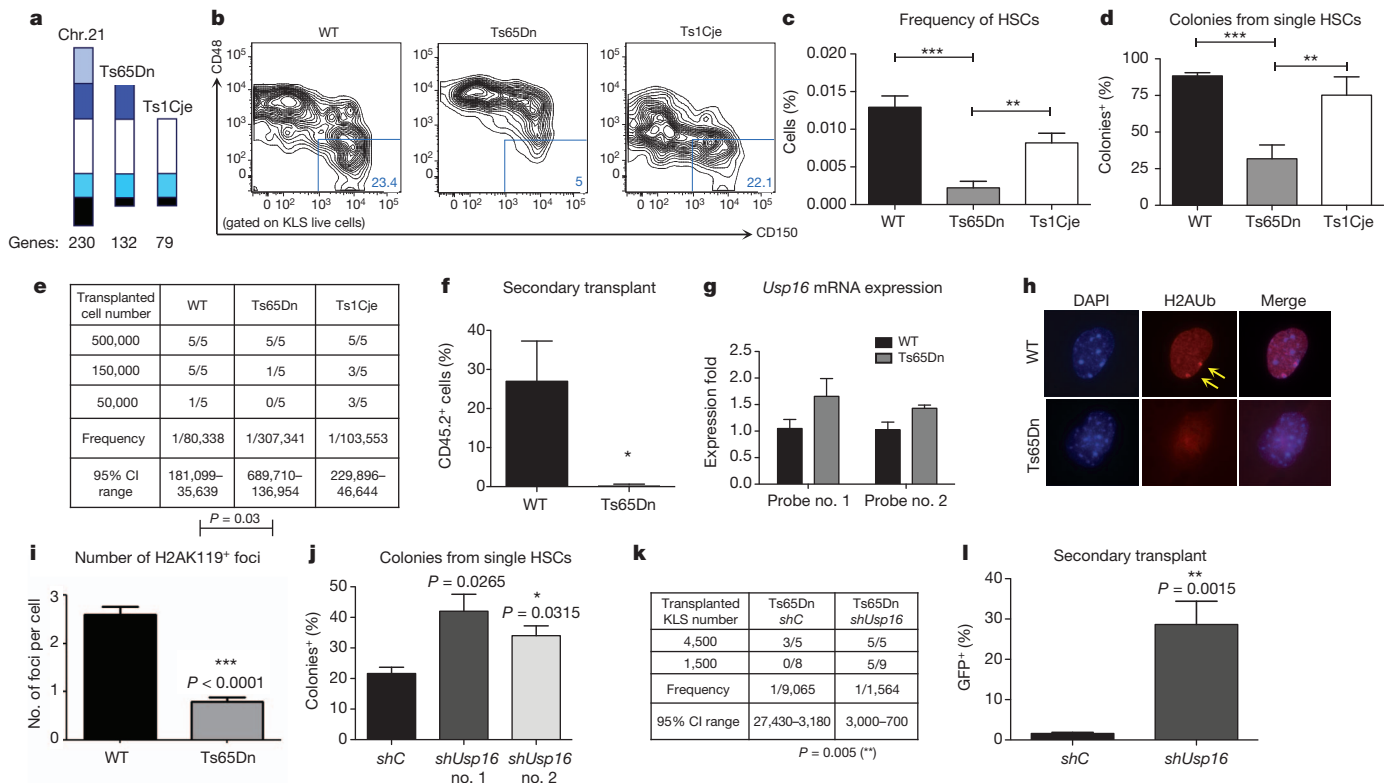
A 50% reduction in *in vitro* colony formation by single CD34<sup>-</sup> CD150<sup>+</sup> CD48<sup>-</sup> KLS cells compared to controls further suggested an impairment of Ts65Dn HSCs (Fig. 1d)<sup>12</sup>. To assess definitively the function of HSCs in Down's syndrome, we performed serial dilution bone marrow transplantations with euploid, Ts65Dn and Ts1Cje cells. Wild-type and Ts1Cje bone marrow cells reconstituted the bone marrow of lethally irradiated recipient mice with a similar HSC calculated frequency (1 out of 80,338 and 1 out of 103,553, respectively). However, there was a threefold reduction in the frequency of Ts65Dn stem cells (1 out of 307,431;  $P = 0.0294$ ) (Fig. 1e and Extended Data Fig. 1c). Furthermore, multi-lineage engraftment of Ts65Dn bone marrow cells was not observed in secondary transplants after 3 months (Fig. 1f and Extended Data Fig. 1d), confirming a severe impairment in the self-renewal ability of Ts65Dn haematopoietic cells.

## Usp16 contributes to HSC defects in mice

*Usp16* is one of the genes uniquely triplicated in Ts65Dn mice but not in Ts1Cje mice. Previous studies have reported that *Usp16* can remove ubiquitin moieties from histone H2A, opposing the ubiquitination function of the Polycomb repressive complex 1 (PRC1)<sup>9</sup>. As Bmi1, a crucial member of the PRC1 complex, is essential for the self-renewal of stem cells in multiple tissues<sup>13–16</sup>, we postulated that an extra copy of *Usp16*, an antagonist of PRC1, could impair the self-renewal ability of stem cells in Ts65Dn mice. Consistent with the presence of three alleles, *Usp16* messenger RNA was expressed 1.5-fold higher in Ts65Dn HSCs compared to control HSCs (Fig. 1g). Moreover, immunofluorescence

<sup>1</sup>Institute for Stem Cell Biology and Regenerative Medicine, Stanford University School of Medicine, Stanford, California 94305, USA. <sup>2</sup>International PhD School in Molecular Medicine, San Raffaele University, Milano, Italy. <sup>3</sup>Program in Cancer Biology, Stanford University School of Medicine, Stanford, California 94305, USA. <sup>4</sup>Department of Neurology, Stanford University School of Medicine, Stanford, California 94305, USA. <sup>5</sup>Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, California 94305, USA. <sup>6</sup>Division of Pediatric Cardiology, Department of Pediatrics, Stanford University School of Medicine, Stanford, California 94305, USA. <sup>7</sup>Division of Pediatric Neurosurgery, Department of Neurosurgery, Stanford University School of Medicine, Stanford, California 94305, USA.





**Figure 1 | *Usp16* contributes to defective HSCs in Ts65Dn mice.** **a**, Down's syndrome mouse models. **b**, **c**, HSC frequency is decreased in Ts65Dn, but not Ts1Cje, bone marrow. Representative plots are shown ( $n = 4$ ). WT, wild type. **d**, Colony-formation assay from single HSCs ( $n = 3$ ). **e**, Limiting dilution analysis (ELDA) of bone marrow cells. **f**, Secondary engraftment 3 months after transplantation, shown as percentage of donor cells in the peripheral blood. Transplants were repeated twice. **g**, Quantification of *Usp16* mRNA.

**h**, **i**, H2AK119 immunofluorescence in HSCs. The arrows indicate the stained H2AK119 foci in one representative picture (**h**). Quantification of positive foci ( $n = 100$  cells analysed in two experiments) (**i**). **j**, Knockdown of *Usp16* in Ts65Dn HSCs rescues colony formation ( $n = 3$ ). **k**, **l**, Knockdown of *Usp16* improves the engraftment potential of Ts65Dn and their ability to undergo secondary transplantation. Transplants were repeated twice. 95% confidence interval (CI) is shown as the interval between lower and upper values.

studies showed a twofold reduction in the number of ubiquitinated chromatin foci in Ts65Dn CD34<sup>+</sup> CD150<sup>+</sup> CD48<sup>+</sup> KLS cells (Fig. 1h, i). Both immunofluorescence and western blots analyses also demonstrated a reduction in the levels of H2AK119 ubiquitination in Ts65Dn mouse embryonic fibroblasts (MEFs) (Extended Data Fig. 2).

To determine whether increased levels of *Usp16* contribute to abnormalities of Ts65Dn HSCs, we generated lentiviral constructs encoding green fluorescent protein (GFP) and short interfering RNAs (shRNAs) directed against a scrambled sequence (*shC*) or *Usp16* (*shUsp16* no. 1 and *shUsp16* no. 2). These hairpins reduce *Usp16* expression to 40–50%, leading to a final expression level similar to the one observed in control animals (Extended Data Fig. 3a). In total, 40% of single CD34<sup>+</sup> CD150<sup>+</sup> CD48<sup>+</sup> KLS Ts65Dn bone marrow cells infected with the *shUsp16* lentivirus formed *in vitro* colonies in MethoCult, compared to 20% of the *shC*-infected cells (Fig. 1j) ( $P < 0.04$ ).

To assess the effect of *Usp16* *in vivo*, Ts65Dn KLS cells were infected with *shUsp16* or *shC* lentivirus and then injected into lethally irradiated C57BL/6 recipient mice. *shUsp16*-infected cells, but not *shC*-infected cells, significantly engrafted in recipient mice (Fig. 1k and Extended Data Fig. 3b). Unlike Ts65Dn bone marrow cells treated with a control lentivirus, the cells transduced with the *shUsp16* lentivirus could also give rise to multipotent differentiation upon serial transplantation (Fig. 1l and Extended Data Fig. 3c). These results demonstrate that the self-renewal defect of Ts65Dn HSCs can be substantially rescued by downregulation of *Usp16* alone.

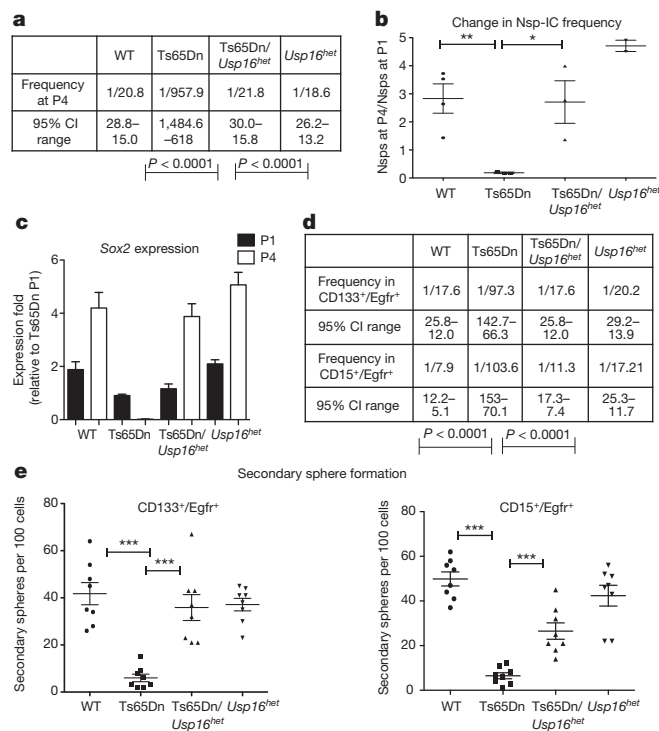
### Role of *Usp16* in Ts65Dn neural progenitor defects

*Usp16* is expressed in a key site of adult neurogenesis, the brain sub-ventricular zone (SVZ) (Extended Data Fig. 4a). Because *Bmi1* is

known to be critical for the maintenance of neural progenitors<sup>15,16</sup>, we proposed that an extra copy of the PRC1 antagonist *Usp16* could have a role in regulating also their expansion in Ts65Dn mice. To test this hypothesis, the SVZ was micro-dissected from 8-week-old mice brains, and Lin<sup>+</sup> cells (CD24<sup>+</sup> CD31<sup>+</sup> CD45<sup>+</sup> Ter119<sup>+</sup>) were enriched by flow cytometry and assayed for neurosphere formation during serial passaging *in vitro*. As reported previously, *in vitro* sphere formation by Ts65Dn neural progenitors was impaired<sup>7,8,17</sup>. By the fourth passage, 1 out of 21 wild-type cells formed neurospheres, whereas only 1 out of 958 Ts65Dn Lin<sup>+</sup> cells could do so ( $P < 0.0001$ ) (Fig. 2a and Extended Data Fig. 4b). To define the role of an extra copy of *Usp16* in neural progenitor expansion, Ts65Dn mice were bred with mice in which one of the normal *Usp16* alleles was mutated (*Usp16*<sup>het</sup>). The offspring had a normal diploid dosage of *Usp16*, but retained three copies of the other genes present in the Ts65Dn parental strain (Ts65Dn/*Usp16*<sup>het</sup> mice). Notably, Ts65Dn/*Usp16*<sup>het</sup> neural progenitor cells maintained a normal frequency of neurosphere-initiating cells (Nsp-ICs) during serial passage ( $P < 0.0001$ ) (Fig. 2a, b).

After four passages, expression of Sox2 (a known neural progenitor cell marker)<sup>18</sup> increased twofold in control Lin<sup>+</sup> neurospheres but was not detected in Ts65Dn Lin<sup>+</sup> neurospheres. By contrast, Ts65Dn/*Usp16*<sup>het</sup> cells express comparable levels of Sox2 to control cells (Fig. 2c).

Next we analysed CD133<sup>+</sup> Egfr<sup>+</sup> Lin<sup>+</sup> and CD15<sup>+</sup> Egfr<sup>+</sup> Lin<sup>+</sup> SVZ cells<sup>19,20</sup>, which are enriched for neural progenitor cells. These cells express high levels of nestin and Sox2 and retain the ability to expand *in vitro* upon serial passaging (Extended Data Fig. 5). Sphere limiting dilution analyses of either CD133<sup>+</sup> Egfr<sup>+</sup> Lin<sup>+</sup> cells or CD15<sup>+</sup> Egfr<sup>+</sup> Lin<sup>+</sup> cells revealed a significant decrease in the frequency of Nsp-ICs in the Ts65Dn mice compared to wild-type or Ts65Dn/*Usp16*<sup>het</sup> mice

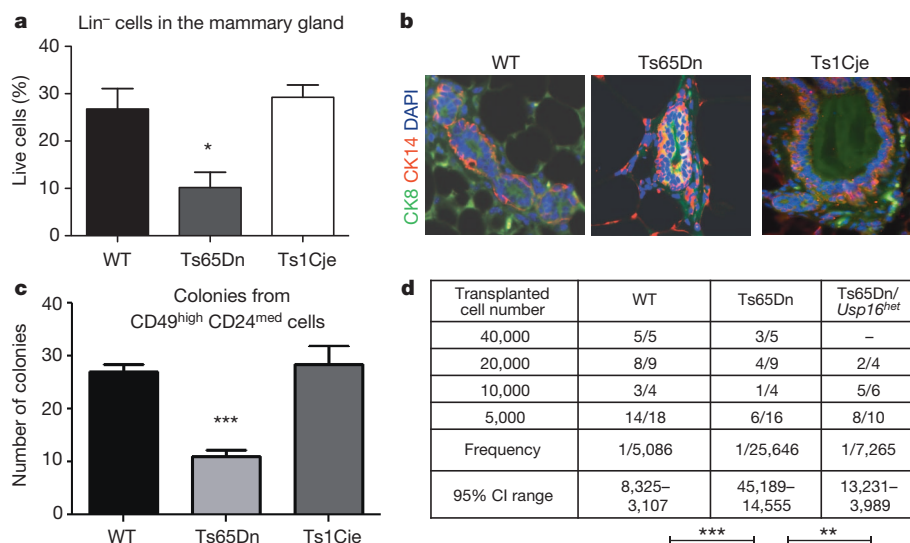


**Figure 2 | Ts65Dn mice, but not Ts65Dn/*Usp16*<sup>het</sup> mice, show defective neural progenitor cells.** **a**, ELDA for SVZ Lin<sup>+</sup> cells (at passage 4 (P4)). **b**, **c**, Nsp-IC frequency (**b**) and Sox2 mRNA levels (**c**) decrease only in Ts65Dn culture. **d**, **e**, ELDA and secondary sphere formation experiments with the indicated populations show a neural progenitor expansion deficit specific to Ts65Dn mice.  $n = 3$  per group in all experiments.

(Fig. 2d and Extended Data Fig. 4c). Moreover, Ts65Dn, but not Ts65Dn/*Usp16*<sup>het</sup>, neural progenitors showed a significant decrease in formation of secondary neurospheres ( $P < 0.0001$ ) (Fig. 2e). Taken together, these data show that Ts65Dn mice have a defect in the expansion of neural progenitors that can be partially rescued by eliminating an extra allele of *Usp16*.

### Role of *Usp16* in mammary epithelium of Ts65Dn mice

Loss of Bmi1 results in severe mammary epithelium growth defects<sup>14</sup>. We therefore wanted to know whether the extra copy of *Usp16* in Ts65Dn mice might also affect the expansion of mammary epithelial



**Figure 3 | Ts65Dn mammary cells are affected by levels of *Usp16*.**

**a**, Lin<sup>+</sup> cells are reduced in Ts65Dn, but not in Ts1Cje, glands ( $n = 5$ ). **b**, Immunofluorescence staining shows an increase overlapping of luminal (CK8) and basal (CK14) cytokeratins in Ts65Dn glands. **c**, Ts65Dn sorted MRUs form fewer colonies than controls ( $n = 4$ ). **d**, ELDA shows a decreased number of repopulating cells in Ts65Dn, but not in Ts65Dn/*Usp16*<sup>het</sup>, glands. Three independent experiments were performed, two for Ts65Dn/*Usp16*<sup>het</sup>.

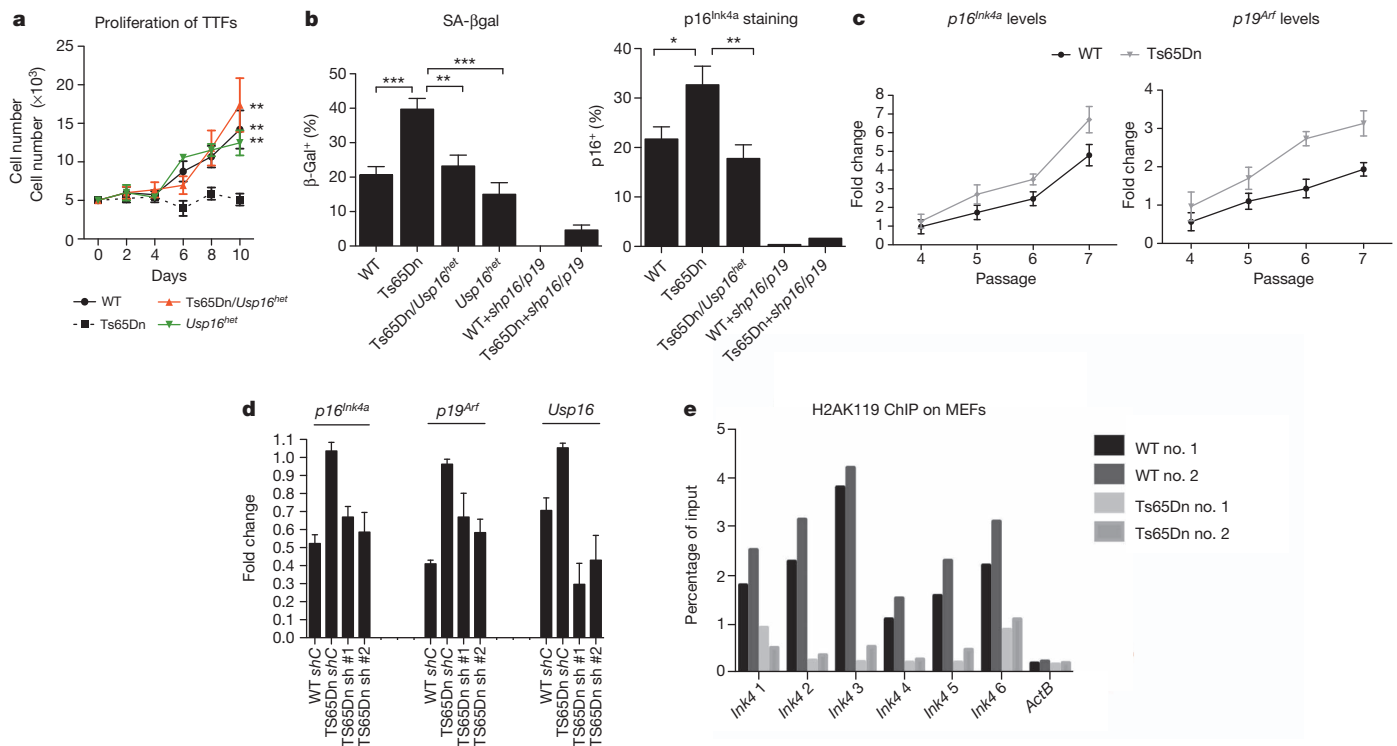
cells. Like HSCs, *Usp16* mRNA expression was increased approximately 1.5-fold in Ts65Dn CD49<sup>+</sup> CD24<sup>medium</sup> Lineage<sup>+</sup> cells (which are enriched for mammary repopulating units (MRUs)<sup>21,22</sup>) compared to control cells (Extended Data Fig. 6a). Moreover, several *Hox* genes, usually repressed by PRC1, were highly expressed in Ts65Dn cells (Extended Data Fig. 6b). Immuno-phenotypic analysis of the breast tissue from Ts65Dn mice revealed a significant reduction of the overall number of CD31<sup>+</sup> CD45<sup>+</sup> Ter119<sup>+</sup> (Lineage<sup>+</sup>) cells (Fig. 3a and Extended Data Fig. 6c). Cytokeratin staining showed that compared to wild-type or Ts1Cje mice, Ts65Dn mice have an increased number of cells that co-express luminal cell cytokeratin-8 and basal cell cytokeratin-14 (Fig. 3b and Extended Data Fig. 6d), suggesting abnormalities in the development of the mammary epithelium.

To functionally test CD49<sup>+</sup> CD24<sup>med</sup> Lin<sup>+</sup> cells, a population enriched with regeneration potential<sup>21,22</sup>, we performed three-dimensional *in vitro* colony assays. Ts65Dn, but not Ts1Cje, cultures showed a reduced number of colonies (Fig. 3c). Moreover, *in vivo* mammary transplantation assays with Lin<sup>+</sup> cells showed a significant decrease in repopulating ability of Ts65Dn Lin<sup>+</sup> cells compared to wild type or Ts65Dn/*Usp16*<sup>het</sup> Lin<sup>+</sup> cells (Fig. 3d).

As in HSCs and neural progenitors, lentiviral downregulation of *Usp16* in Ts65Dn Lin<sup>+</sup> cells partially rescued mammary epithelium defects. There was a twofold increase in the calculated frequency of MRUs in *shUsp16*-infected Lin<sup>+</sup> Ts65Dn cells, and the derived outgrowths were larger compared to cells infected with control shRNA (Extended Data Fig. 6e). However, we were unable to generate mammary glands in secondary transplants of the Ts65Dn/*Usp16*<sup>het</sup> mammary epithelial cells, suggesting that other genes may affect the proliferation of these cells in Ts65Dn mice.

### Modulation of *Ink4a*/*Arf* by *Usp16* in Ts65Dn cells

One of the best-characterized Bmi1/PRC1 target loci is *Cdkn2a*<sup>23</sup>, which encodes two distinct tumour suppressors, p16<sup>Ink4a</sup> and p19<sup>Arf</sup>. Expression of p16<sup>Ink4a</sup> and p19<sup>Arf</sup> normally increases with age in both rodent and human tissues and is involved in senescence-induced loss of proliferation<sup>24–26</sup>. To determine whether *Usp16* trisomy affects proliferation and senescence, we used MEFs and terminal-tip fibroblasts (TTFs). Ts65Dn, but not Ts65Dn/*Usp16*<sup>het</sup>, TTFs showed a marked proliferation deficit and high levels of senescence as shown by senescence-associated  $\beta$ -galactosidase staining and by p16<sup>Ink4a</sup> expression (Fig. 4a, b). These defects were partially rescued by a shRNA targeting *Cdkn2a* (Fig. 4b and Extended Data Fig. 7a–c). Ts65Dn MEF cultures showed high levels of senescence and a rapid increase of p16<sup>Ink4a</sup> and p19<sup>Arf</sup> expression levels upon passaging (Fig. 4c and Extended Data Fig. 7d). However, lentiviral downregulation



**Figure 4 | Usp16 contributes to proliferation defects and senescence in Ts65Dn fibroblasts.** **a**, Proliferation of TTFs seeded at P2. **b**, Senescent cells are frequent in Ts65Dn, but not in Ts65Dn/Usp16<sup>het</sup>. TTFs as shown by senescence-associated  $\beta$ -galactosidase (SA- $\beta$ gal) and p16<sup>Ink4a</sup> staining at P3. Downregulation of *Cdkn2a* blocks senescence. **c**, p16<sup>Ink4a</sup> and p19<sup>Arf</sup>

mRNA levels increase rapidly during passaging in Ts65Dn MEFs.

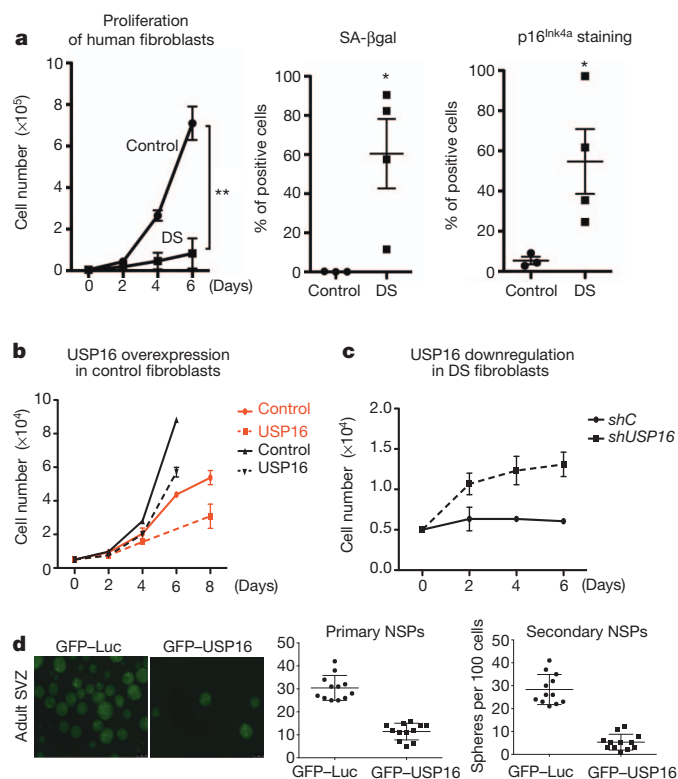
**d**, Downregulation of *Usp16* normalizes mRNA expression of *Ink4a/Arf* by P6 Ts65Dn MEFs. **e**, Chromatin immunoprecipitation (ChIP) analyses at the *Cdkn2a* locus show lower levels of H2AK119 binding by two Ts65Dn MEF chromatin samples.

of *Usp16* with two different hairpins decreased p16<sup>Ink4a</sup> and p19<sup>Arf</sup> expression (Fig. 4d). Western blots showed decreased levels of H2AK119 ubiquitination in Ts65Dn MEFs (Extended Data Fig. 2)

and chromatin immunoprecipitation analyses demonstrated a decrease in H2AK119 ubiquitin on the *Ink4a/Arf* locus in Ts65Dn MEFs (Fig. 4e). These results strongly suggest that trisomy of *Usp16* contributes to the cellular defects observed in Ts65Dn mice through increased removal of ubiquitin moieties from H2A, with an impact on the *Cdkn2a* site. However, mutational analyses of the deubiquitinase domains would be required to confirm this model.

### Potential role USP16 in human Down's syndrome

As reported previously<sup>27,28</sup>, fibroblasts derived from Down's syndrome patients showed a strong proliferation defect. Moreover, they appear to senesce prematurely (Fig. 5a). Gain-of-function experiments were performed using normal fibroblasts to determine whether USP16 overexpression might contribute to the proliferation defects. Transduction of normal foreskin fibroblasts with a vector overexpressing USP16 slowed their proliferation (Fig. 5b and Extended Data Fig. 8a). Conversely, shRNA-mediated downregulation of USP16 or overexpression of BMI1 resulted in increased proliferation capacity of Down's syndrome fibroblasts (Fig. 5c and Extended Data Fig. 8b, c). Interestingly, overexpression of USP16 in two different cultures of human neural progenitor cells (from one paediatric and one adult patient) reduced their *in vitro* expansion potential and the formation of neurospheres



**Figure 5 | Human cells are affected by USP16 levels.** **a**, Proliferation analysis, as well as SA- $\beta$ gal and p16<sup>Ink4a</sup> staining, of three control and four Down's syndrome (DS) human fibroblast cultures show growth impairment and senescence of Down's syndrome cells. **b**, **c**, Lentiviral-induced overexpression of USP16 decreases the proliferation of two different control fibroblast lines (**b**), whereas downregulation of USP16 in Down's syndrome fibroblasts promotes proliferation (**c**). **d**, Overexpression of USP16 reduces the formation of neurospheres derived from human adult SVZ cells. The right panel quantifies the number of spheres in the first and second passages.  $P < 0.0001$ . All the experiments were replicated at least twice. Luc, luciferase.



(Fig. 5d and Extended Data Fig. 8d). These experiments suggest a role for the third copy of *USP16* in the pathologies associated with Down's syndrome.

## Discussion

Our data show that the change in self-renewal capacity of HSCs and neural progenitors, as well as the proliferation defects of mammary epithelial cells and fibroblasts in Ts65Dn mice, is linked in part to trisomy of *Usp16*, a negative modifier of the PRC1 activity<sup>9</sup>. As Bmi1, a component of PRC1, is known to regulate senescence and self-renewal of multiple somatic stem cells, this suggests that some of the pathologies associated with Down's syndrome may result from a stem-cell imbalance secondary to overexpression of *Usp16*. Clearly, other genes also have a role in traits associated with Down's syndrome patients, such as craniofacial abnormalities and perhaps the increased incidence of leukaemia and the decreased incidence of solid tumours (Extended Data Fig. 9)<sup>2,29</sup>, and there are probably differences between humans with Down's syndrome and Ts65Dn mice. Nonetheless, the defects of human fibroblasts mirror those seen in the mouse models, and the effect of overexpression of *Usp16* in human neural progenitors suggest that this gene has a similar function in human Down's syndrome and in Ts65Dn mice. Although we cannot unequivocally state that the HSCs are reduced in Down's syndrome patients, in two children we examined they were not increased as they might be in the liver of some human Down's syndrome fetuses (Extended Data Fig. 9)<sup>30</sup>.

In conclusion, our data demonstrate a new axis of regulation of tissue homeostasis in multiple tissues. We believe that our study has broad implications for understanding one of the most common genetic abnormalities in humans, Down's syndrome, and that targeting the *USP16* pathway may ameliorate some of the associated pathologies.

## METHODS SUMMARY

**Mice.** Nonobese diabetic/severe combined immunodeficiency (NOD/SCID), Ts65Dn, Ts1Cje and euploid littermate mice were purchased from Jackson Laboratories. *Usp16<sup>het</sup>* mice (FVB/N-*Usp16<sup>Tg(Tyr)2414FOve</sup>/Mmjax*) were obtained from the Mutant Mouse Resource Centers (MMRRC).

**Lentivirus.** The lentiviral vector used for downregulation was pSicoR-GFP. *USP16* overexpression vector was obtained by subcloning a *USP16* clone (American Type Culture Collection (ATCC)) into a pCDH-MSCV-GFP vector or an pCDH-EF1A-GFP vector (System Biosciences).

**Human fibroblast.** Human fibroblasts (wild type: CRL-2088, CRL-2076; Down's syndrome: CCL-54, CRL-7090, CRL-7031) were purchased by ATCC. Other fibroblast cultures (Down's syndrome: GM2508 and normal fibroblasts) were kindly provided by J. Weimann and S. Marro.

**Human neurosphere cultures.** Adult SVZ neurospheres were derived from a rapid autopsy. Paediatric SVZ neurospheres were derived during a functional hemispherectomy. Human tissue was acquired under Institutional approved IRB protocol no. 18672 and the Stanford Stem Cell Research Oversight (SCRO) committee, no. 370.

**Statistical analyses.** In all the graphs, bars show average as central values and  $\pm$  s.e.m. as error bars, unless otherwise specified. *P* values were calculated using two-tailed *t*-test analyses. For limiting dilution analyses, ELDA software or L-Calc were used to test inequality between multiple groups. Expected frequencies are reported, as well as the 95% confidence intervals (lower and upper values are indicated). \**P* < 0.05, \*\**P* < 0.01, \*\*\**P* < 0.001

**Online Content** Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 17 October 2012; accepted 5 August 2013.

Published online 11 September 2013.

- Antonarakis, S. E., Lyle, R., Dermitzakis, E. T., Reymond, A. & Deutsch, S. Chromosome 21 and Down syndrome: from genomics to pathophysiology. *Nature Rev. Genet.* **5**, 725–738 (2004).
- Yang, Q., Rasmussen, S. A. & Friedman, J. M. Mortality associated with Down's syndrome in the USA from 1983 to 1997: a population-based study. *Lancet* **359**, 1019–1025 (2002).
- Roth, G. M., Sun, B., Greensite, F. S., Lott, I. T. & Dietrich, R. B. Premature aging in persons with Down syndrome: MR findings. *Am. J. Neuroradiol.* **17**, 1283–1289 (1996).

- Zigman, W. B. & Lott, I. T. Alzheimer's disease in Down syndrome: neurobiology and risk. *Ment. Retard. Dev. Disabil. Res. Rev.* **13**, 237–246 (2007).
- Liu, L. & Rando, T. A. Manifestations and mechanisms of stem cell aging. *J. Cell Biol.* **193**, 257–266 (2011).
- Reeves, R. H. *et al.* A mouse model for Down syndrome exhibits learning and behaviour deficits. *Nature Genet.* **11**, 177–184 (1995).
- Sago, H. *et al.* Genetic dissection of region associated with behavioral abnormalities in mouse models for Down syndrome. *Pediatr. Res.* **48**, 606–613 (2000).
- Sago, H. *et al.* Ts1Cje, a partial trisomy 16 mouse model for Down syndrome, exhibits learning and behavioral abnormalities. *Proc. Natl Acad. Sci. USA* **95**, 6256–6261 (1998).
- Joo, H.-Y. *et al.* Regulation of cell cycle progression and gene expression by H2A deubiquitination. *Nature* **449**, 1068–1072 (2007).
- Lorenzo, L. P. E. *et al.* Defective hematopoietic stem cell and lymphoid progenitor development in the Ts65Dn mouse model of Down syndrome: potential role of oxidative stress. *Antioxid. Redox Signal.* **15**, 2083–2094 (2011).
- Chao, M. P., Seita, J. & Weissman, I. L. Establishment of a normal hematopoietic and leukemia stem cell hierarchy. *Cold Spring Harb. Symp. Quant. Biol.* **73**, 439–449 (2008).
- Akala, O. O. *et al.* Long-term haematopoietic reconstitution by *Trp53<sup>-/-</sup>p16<sup>ink4a</sup><sup>-/-</sup>p19<sup>Arf</sup><sup>-/-</sup>* multipotent progenitors. *Nature* **453**, 228–232 (2008).
- Park, I.-K. *et al.* Bmi-1 is required for maintenance of adult self-renewing haematopoietic stem cells. *Nature* **423**, 302–305 (2003).
- Pietersen, A. M. *et al.* Bmi-1 regulates stem cells and proliferation and differentiation of committed cells in mammary epithelium. *Curr. Biol.* **18**, 1094–1099 (2008).
- Molofsky, A. V. *et al.* Bmi-1 dependence distinguishes neural stem cell self-renewal from progenitor proliferation. *Nature* **425**, 962–967 (2003).
- Cao, G. *et al.* Bmi-1 absence causes premature brain degeneration. *PLoS ONE* **7**, e32015 (2012).
- Lorenzi, H. A. & Reeves, R. H. Hippocampal hypocellularity in the Ts65Dn mouse originates early in development. *Brain Res.* **1104**, 153–159 (2006).
- Ellis, P. *et al.* SOX2, a persistent marker for multipotential neural stem cells derived from embryonic stem cells, the embryo or the adult. *Dev. Neurosci.* **26**, 148–165 (2004).
- Fischer, J. *et al.* Prospective isolation of adult neural stem cells from the mouse subependymal zone. *Nature Protocols* **6**, 1981–1989 (2011).
- Pastrana, E., Cheng, L.-C. & Doetsch, F. Simultaneous prospective purification of adult subventricular zone neural stem cells and their progeny. *Proc. Natl Acad. Sci. USA* **106**, 6387–6392 (2009).
- Stingl, J. *et al.* Purification and unique properties of mammary epithelial stem cells. *Nature* **439**, 993–997 (2006).
- Shackleton, M. *et al.* Generation of a functional mammary gland from a single stem cell. *Nature* **439**, 84–88 (2006).
- Jacobs, J. J., Kieboom, K., Marino, S., DePino, R. A. & van Lohuizen, M. The oncogene and Polycomb-group gene *bmi-1* regulates cell proliferation and senescence through the *ink4a* locus. *Nature* **397**, 164–168 (1999).
- Liu, Y. *et al.* Expression of *p16<sup>INK4a</sup>* in peripheral blood T-cells is a biomarker of human aging. *Aging Cell* **8**, 439–448 (2009).
- Krishnamurthy, J. *et al.* *Ink4a/Arf* expression is a biomarker of aging. *J. Clin. Invest.* **114**, 1299–1307 (2004).
- Janzen, V. *et al.* Stem-cell ageing modified by the cyclin-dependent kinase inhibitor *p16<sup>INK4a</sup>*. *Nature* **443**, 421–426 (2006).
- Kimura, M. *et al.* Proliferation dynamics in cultured skin fibroblasts from Down syndrome subjects. *Free Radic. Biol. Med.* **39**, 374–380 (2005).
- Carmeliet, G., David, G. & Cassiman, J. J. Cellular ageing of Alzheimer's disease and Down syndrome cells in culture. *Mutat. Res.* **256**, 221–231 (1991).
- Arron, J. R. *et al.* NFAT dysregulation by increased dosage of *DSCR1* and *DYRK1A* on chromosome 21. *Nature* **441**, 595–600 (2006).
- Roy, A. *et al.* Perturbation of fetal liver hematopoietic stem and progenitor cell development by trisomy 21. *Proc. Natl Acad. Sci. USA* **109**, 17579–17584 (2012).

**Acknowledgements** We thank S. Marro for help in establishing MEF and TTF lines; I. L. Weissman for sharing C57BL/6 CD45.1 mice; the animal core facility of SIM1, in particular A. Valledifera and M. Alvarez. This study was supported in part by a California Institute for Regenerative Medicine (CIRM) basic biology award III, National Institutes of Health grant numbers CA100225 and CA154209, CIRM bridges fellowship (V.H.-A.), Fondazione Umberto Veronesi (B.N.d.R.), Breast Cancer Research Program sponsored by Department of Defense (S.S.), The Breast Cancer Research Foundation (BCRF), Stanford Graduate Fellowship (Y.O.) and the Down Syndrome Research and Treatment Foundation and the Fidelity Foundation (C.C.G.).

**Author Contributions** M.A. and M.F.C. designed the project and wrote the manuscript. M.A. identified *Usp16* in a genetic screen, constructed the lentiviral constructs and carried out bone marrow experiments. S.S. performed most of the breast experiments. S.S.M. performed most of the neural experiments. M.A., B.N.d.R., D.Q., A.K., V.H.-A., J.R. and Y.O. contributed to mouse colony handling, lentivirus preparation and mouse fibroblast experiments. M.Q. performed the imaging studies. V.M.R. provided the bone marrow specimens. S.C. and C.C.G. oversaw the neural progenitor experiments, provided animals and reagents and revised the manuscript. Project management was carried out by M.A.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.F.C. ([mfcclarke@stanford.edu](mailto:mfcclarke@stanford.edu)).

## METHODS

**Statistical analyses.** In all the graphs, bars show average as central values and  $\pm$  s.e.m. as error bars, unless otherwise specified. *P* values were calculated using two-tailed *t*-test analyses. For limiting dilution analyses, ELDA software or L-Calculator were used to test inequality between multiple groups. Expected frequencies are reported, as well as the 95% confidence intervals (lower and upper values are indicated). \**P* < 0.05, \*\**P* < 0.01, \*\*\**P* < 0.001

**Mice.** Ts65Dn, Ts1Cje and euploid littermates mice (mixed background B6EiC3SnF1/J) were purchased from Jackson Laboratories. Mice were genotyped by real-time PCR or by traditional PCR as described previously<sup>31</sup> or according to animal's provider. Control littermates were used as wild-type mice. These mice are heterozygous for B6 and C3H alleles at all loci in their genome. C57BL/6 CD45.1 mice, used for bone marrow transplantation experiments, were kindly provided by I. Weissman. *Usp16<sup>het</sup>* mice (FVB/N-*Usp16<sup>Tg(Tyr)2414FOve</sup>/Mmjax*) were ordered from Mutant Mouse Regional Resource Centers (MMRRC). Weaning-age nonobese diabetic/severe combined immunodeficiency (NOD/SCID) female mice were purchased by Jackson Laboratories. Mice were housed in accordance with the guidelines of Institutional Animal Care Use Committee.

**Bone marrow and peripheral blood analyses.** Isolation and analyses of bone marrow cells were performed as described previously<sup>12</sup>. In brief, bone marrow cells were isolated by crushing long bones and hips with mortar and pestle in calcium- and magnesium-free Hank's balanced salt solution (HBSS) with 2% heat-inactivated bovine serum. The cells were drawn by passing through a 25G needle several times, treated with ammonium chloride potassium (ACK) for 1 min and filtered with a 40-mm nylon mesh. Before sorting, progenitor cells were enriched through magnetic isolation with mouse lineage cell depletion kit (Miltenyi Biotec) using an autoMACS pro Separator. Markers used for identification of lineage cells were CD3, CD5, CD8, Gr-1, B220 and Ter119.

For peripheral blood analysis, red blood cells were lysed with hypotonic buffer, and nucleated cells were stained with antibodies against CD45.1, CD45.2, Ter119, Gr-1, Mac-1, CD3 and B220.

Details of antibodies used for analyses and sorting of hematopoietic cells: c-Kit APC780, CD45R (B220) phycoerythrin (PE)-Cy5, CD45R (B220) PE, Ly6G (Gr-1) APC, CD8a PE-Cy5, CD3e Cy5, CD3e PE, CD135 (Flt3) APC, CD34 FITC, CD45.1 PE, CD45.2 FITC (eBioscience); CD150 PE, CD48 APC, Sca1 PacBlue, Ter119 PacBlue, Ter119 PE-Cy5, CD5 PE-Cy5, CD11b (Mac1) APC (Biolegend).

Cells were gated on the basis of forward- and side-scatter profiles, and live/dead discrimination was obtained with 7-amino-actinomycin D (7-AAD) or DAPI. Analyses and sorting was performed using a FACS Aria II (BD Biosciences). Data were also analysed with FlowJo software.

**Bone marrow transplants.** Recipient C57BL/6 CD45.1 mice (8–12 weeks old) were lethally irradiated (1,140 rad), with two doses of radiations delivered 3 h apart. Bone marrow single-cell suspensions were obtained from long bones and hips of Ts65Dn, Ts1Cje and wild-type mice (8–12 weeks old) and treated for 1 min with ACK for red-cell lysis. Only donor animals with a matching haplotype for the major histocompatibility antigens were used (*H-2K<sup>b/b</sup>*)<sup>10</sup>. Recipient mice were competitively reconstituted by retro-orbital venous sinus injection with three different doses ( $5 \times 10^5$  cells,  $1.5 \times 10^5$  cells,  $0.5 \times 10^5$  cells) of whole bone marrow cells from donor mice and a radioprotective dose of  $3 \times 10^5$  bone marrow cells from non-irradiated C57BL/6/Ka-CD45.1 mice. Five mice were used for each group. Mice were analysed monthly for reconstitution of peripheral blood.

For secondary transplants,  $5 \times 10^6$  bone marrow cells were collected at least 4 months after primary transplants from engrafted recipients, and injected in lethally irradiated C57BL/6 CD45.1 mice. Reconstitution was measured in peripheral blood at the indicated time points.

For transplantation of lentivirus-infected Ts65Dn marrow cells, KLS cells were isolated and sorted from 8–12-week-old Ts65Dn *H-2K<sup>b/b</sup>* mice. Sorted KLS were incubated overnight with the indicated lentivirus (multiplicity of infection (m.o.i.) = 200). The morning after, KLS were washed, mixed with a radioprotective dose of  $3 \times 10^5$  bone marrow cells from non-irradiated C57BL/6/Ka-CD45.1 mice, and injected in lethally irradiated mice. In parallel, 48 h after infection we verified by FACS the level of GFP expression in infected cells, to make sure that the level of lentiviral integration would be similar between multiple samples.

For all transplantation experiments, we considered as repopulated by donor cells the mice that had more than 1% donor-derived (CD45.2<sup>+</sup>) cells in both lymphoid (CD3<sup>+</sup> and B220<sup>+</sup>) and myeloid (Gr-1<sup>+</sup> and Mac-1<sup>+</sup>) subpopulations. Frequency of long-term reconstituting cells from limiting dilution experiments was calculated using ELDA software<sup>32</sup>. All the transplants were repeated at least twice.

**In vitro colony formation of HSCs.** For MethoCult cultures, we double-sorted single HSCs into U-bottom 96-well plates containing 100  $\mu$ l of MethoCult GF M3434 media (Stemcell Technologies) as described previously<sup>12</sup>. Positive colonies were scored after 7 days of culture.

**Neurosphere assays.** Wild-type, Ts65Dn, Ts65Dn/*Usp16<sup>het</sup>* and *Usp16<sup>het</sup>* mice were euthanized by CO<sub>2</sub>, decapitated, and their brains immediately removed. The SVZ was micro-dissected and stored in ice-cold PBS for further processing. The tissue was digested using TryPLE express (Invitrogen) and DNase I (250 U ml<sup>-1</sup>) at 37 °C for 10 min followed by trituration using a fire-polished pipette. Digested tissue was washed in ice-cold PBS without calcium and magnesium, filtered through a 40- $\mu$ m filter and re-suspended in neurospheres growth media that is, Neurobasal-A (Invitrogen) and Dulbecco's modified Eagle's medium (DMEM) F/12 (1:1) supplemented with 2% B27-A (Invitrogen), 1% N2 (Invitrogen), mouse recombinant epidermal growth factor (EGF; 20 ng ml<sup>-1</sup>) and basic fibroblast growth factor (bFGF; 20 ng ml<sup>-1</sup>) (Shenandoah Biotechnology) and 2  $\mu$ g ml<sup>-1</sup> heparin (Sigma). Lineage cells were depleted using mouse CD45, CD31, CD34 and Ter119 microbeads (Miltenyi Biotec) and the negative fraction collected. For FACS analysis the cells were stained with anti-CD15–fluorescein isothiocyanate (FITC) (MMA; BD), anti-prominin1–APC (eBiosciences) and biotinylated EGF complexed with PE-Cy7–streptavidin (2  $\mu$ g ml<sup>-1</sup>; Invitrogen).

For limiting dilution analysis, cells were directly plated in 96 wells in limiting dilution down to one cell per well. Each plating dose was done in 24 wells and the number of wells with neurospheres was counted after 10 days. For serial passaging, neurospheres from each passage were collected and re-plated either as 100 cells per well in a 48-well dish or in a similar limiting dilution fashion as described above.

Micro-dissected mouse SVZ was dissociated as described above and FACS-sorted into Lin<sup>-</sup>CD15<sup>+</sup>Egfr<sup>+</sup>CD24<sup>-</sup> and Lin<sup>-</sup>CD15<sup>-</sup>Egfr<sup>-</sup>CD24<sup>-</sup>, or Lin<sup>-</sup>Prom1<sup>+</sup>Egfr<sup>+</sup>CD24<sup>-</sup> and Lin<sup>-</sup>Prom1<sup>-</sup>Egfr<sup>-</sup>CD24<sup>-</sup> cells. In total, 10,000 cells were sorted for each population and concentrated by centrifugation. Cells were collected on a glass slide by cytospin and air dried before methanol fixation. Fixed cells were blocked using 5% normal goat serum and incubated in rabbit anti-Sox2 antibody (ab97959, Abcam) and mouse anti-nestin antibody (2Q178, Abcam), followed by secondary anti-mouse Alexa 594 and anti-rabbit Alexa 488, respectively. Fluorescence was visualized using standard epifluorescence microscopy. Six random fields were imaged and data represented as ratio of Sox2- or nestin-positive cells per DAPI-positive cells in the same field.

**Mouse breast analyses.** Mammary glands were dissected from either wild-type, Ts65Dn or Ts1Cje mice and analysed as described previously<sup>21</sup>. In brief, the glands were digested in collagenase/hyaluronidase followed by ACK lysis, trypsin and DNAase/dispase. The cells were then stained with the following antibodies: CD49f APC, CD24 PE-Cy7, CD45 PacBlue, Ter119 PacBlue, CD31 PacBlue (Biolegend).

Cells were gated on the basis of forward- and side-scatter profiles, and live/dead discrimination was obtained with DAPI. Analyses and sorting was performed using a FACS Aria II (BD Biosciences).

**In vitro mammary colony-forming assays.** 96-well ultra-low attachment plates (BD Biosciences) were prepared with a feeder layer of irradiated L-WNT3a mixed with 60  $\mu$ l of growth factor reduced Matrigel (BD) per well. 1000 sorted MRUs from WT, Ts65Dn or Ts1Cje mice were then plated into liquid media as previously described<sup>33,34</sup>. 10% FBS and 2.5% growth factor reduced Matrigel were added as supplements.

**Mammary transplants.** Lin<sup>-</sup> (CD45<sup>-</sup>CD31<sup>-</sup>Ter119<sup>-</sup>) cell populations were isolated from 12-week-old mice in staining media and re-suspended in 10  $\mu$ l of sterile PBS + 30% Matrigel per transplant before being injected into the cleared fat pads of 21–28-day-old recipient NOD/SCID mice as described previously<sup>21</sup>. All transplants were allowed to grow for at least 6 weeks but not more than 10 weeks before analysis. For knockdown of *Usp16*, Lin<sup>-</sup> cells were infected with either control lentivirus or shRNA against *Usp16* in DMEM/F12 plus 10% FBS overnight. The cells were then washed and re-suspended in sterile PBS plus 30% Matrigel for transplant.

For mammary transplant outgrowth area calculation, NIH Image J software was used. In brief, GFP-positive mammary ducts were measured with the free-hand tool by drawing a shape around the duct. Measurements were performed in a 'blind' fashion and at the same magnification for all samples. The entire fat pad was used to determine the maximum area coverage (100%). Only positive outgrowths were used in the measurement.

**Immunofluorescence of mammary tissue.** 12-week-old mice were euthanized and mammary glands were surgically removed. Glands were fixed in formalin overnight and then transferred to 70% ethanol. They were then embedded in paraffin and sectioned for histology. For staining the slides were de-paraffinized in xylene and alcohol grades. Antigen retrieval was carried out in Tris-EDTA buffer by heating in a microwave for 20 min. Primary antibodies CK14 AF64 (Covance) and CK8 (Developmental Studies Hybridoma Bank, University of Iowa) were applied overnight. Secondary antibodies were anti-rat DyLight 488 and anti-rabbit DyLight 594 (both from Jackson Laboratories). Sections were then mounted using ProLong Antifade reagent (Invitrogen). Images were taken with a NIKON inverted



microscope. Lumosity software was used to quantify the Pearson's correlation of overlapping between CK8 and CK14.

**Western blot and ChIP.** For western blot analyses, chromatin extracts were prepared with subcellular protein fractionation kit (Thermo Scientific). H2AK119 antibody (rabbit) was purchased from Cell Signaling; H2A antibody was purchased from Millipore.

ChIP was performed essentially as described previously<sup>35</sup> using the polyclonal antibody for H2AK119 (Cell Signaling).

**Lentivirus preparation.** The lentiviral vector that we used for downregulation was vector pSicoR-GFP<sup>36</sup>. We cloned the following hairpins: *shC* (5'-TTCTCCG AACGTGTACAGT-3') *shUsp16* no. 1 (5'-CGAGTGTCTGTATTCCTTATAT-3'), *shUsp16* no. 2 (5'-TTCTCTGGAAATACACCTATG-3'), *shp16* (5'-CATCAAGA CATCGTGCATAT-3'), *shp19* (5'-GCCATCTAAACGGTTCAGTTT-3'), human *shUSP16* (5'-TATATCAGTTCACCCGTAAT-3'). A lentivirus construct expressing mCherry and Bmi1 (pEIZ-HIV-mCherry-Bmi1) was kindly provided by Y. Shimono<sup>37</sup>. USP16 overexpression vector was obtained by subcloning *USP16* clone (ATCC) in pCDH-MSCV-GFP vector or pCDH-EF1A-GFP vector (System Biosciences).

Viruses were produced in 293T cells with a second-generation lentivirus system. Supernatants were collected at 48 h and 72 h, and concentrated through ultracentrifugation<sup>38</sup>. Viral titres were calculated by FACS analyses of 293T cells infected with serial dilution of concentrated virus.

**MEFs, TTFs and human fibroblasts.** MEFs were generated from E14.5 embryos obtained from Ts65Dn mothers. Genotype was verified by real-time PCR. Cells were passaged 1:4 when almost confluent. To culture mouse primary TTFs, the skin was peeled from tail tips of wild-type ( $n = 4$ ), Ts65Dn ( $n = 3$ ), Ts65Dn/*Usp16*<sup>het</sup> ( $n = 3$ ) and *Usp16*<sup>het</sup> ( $n = 1$ ) 8-week-old mice, finely minced with a blade and shortly digested in trypsin. Resulting pieces were incubated in DMEM containing 20% FBS for 3 weeks. Derived fibroblasts were then passaged to new plates and considered P2. Human fibroblasts (wild type: CRL-2088, CRL-2076; Down's syndrome: CCL-54, CRL-7090, CRL-7031) were purchased by ATCC. Other fibroblast cultures (Down's syndrome: GM2508 and normal fibroblasts) were kindly provided by J. Weimann and S. Marro (Stanford University).

**Fibroblast proliferation, SA- $\beta$ gal and p16 staining.** In total,  $5 \times 10^3$  fibroblasts were seeded in a 24-well plate and viable cells were counted by trypan blue exclusion at the indicated time points.

For SA- $\beta$ gal staining of senescent cells, the Senescence Detection Kit (Abcam, ab65351) was used according to manufacturer's protocol.

For p16 staining, fibroblasts were permeabilized with 0.2% Triton-PBS, blocked in 3% BSA-PBS and stained with a mouse anti-human p16 (JC8, Santa Cruz Biotechnology) or rabbit anti-mouse p16. Specific secondary antibodies (Alexa Fluor 488 anti-mouse and Alexa Fluor 647 anti-rabbit, Invitrogen) were used (1:1,000).

SA- $\beta$ Gal and p16 staining were detected by microscope at  $\times 10$  and the positive cells were evaluated in three different fields per well. Three technical replicates were performed.

In experiments with infected cells, cells were sorted on the basis of GFP or mCherry expression and expression of the construct was verified by real-time PCR.

**Human neurosphere assays.** Human adult SVZ neurospheres were derived from a rapid autopsy excised adult human brain (National Disease Research Interchange) SVZ. Paediatric SVZ neurospheres were derived from subventricular tissue surgically excised during a functional hemispherectomy. Human tissue was acquired under institutional-approved IRB protocol no. 18672 and the Stanford Stem Cell Research Oversight (SCRO) committee no. 370.

Human tissue was digested using collagenase IV (1 mg ml<sup>-1</sup>) and DNase I (250 U ml<sup>-1</sup>) at 37 °C for 30 min followed by trituration using a fire-polished pipette. Digested tissue was washed in ice-cold HBSS without calcium and magnesium, filtered through a 40- $\mu$ m filter and re-suspended in neurospheres growth media that is, Neurobasal-A (Invitrogen) and DMEM F/12 (1:1) supplemented with 2% B27-A (Invitrogen), 1% N2 (Invitrogen), human recombinant EGF (20 ng ml<sup>-1</sup>) and bFGF (20 ng ml<sup>-1</sup>) (Shenandoah Biotechnology) LIF (20 ng ml<sup>-1</sup>) (millipore) and 2  $\mu$ g ml<sup>-1</sup> heparin (Sigma).

For lentiviral infection, neurospheres were dissociated into single cells and 50,000 cells were infected with pCDH-EF1a-GFP-Luc or pCDH-EF1a-GFP-USP16 at a m.o.i. of 25 and followed for GFP expression. Live GFP<sup>+</sup> cells were clone sorted at a count of 100 cells per well were in a 96-well plate and number of neurospheres was counted after 7 days in culture. For secondary sphere forming ability, neurospheres formed during the previous experiment were re-dissociated and again plated at 100 cells per well in a 96 well-plate and number of spheres per well were counted. The experiment was repeated at two different passages.

**RNA expression analyses.** For real-time analyses, cells were collected in Trizol (Invitrogen), and RNA was extracted following the manufacturer's protocol. Complementary DNA was obtained using Superscript III First Strand Synthesis (Invitrogen).

Real-time reactions were assembled using Taqman probes (Applied Biosystems) in accordance with the manufacturer's directions. Expression data were normalized by the expression of housekeeping genes *ActB* and *Gapdh*. Probes used in this study: *Usp16* (mouse: Mm\_00470393, Mm\_00470406; human: Hs01062190\_m1, Hs01062191\_m1), *p16*<sup>INK4a</sup> (Mm\_01257348, Mm\_00494449), *p19*<sup>Arf</sup> (Mm\_00486943), *Actb* (Mm\_00607939), *Gapdh* (Mm\_99999915), *Hoxa1* (Mm00439359\_m1), *Hoxa3* (Mm01326402\_m1), *Hoxa5* (Mm01326402\_m1), *Bmi1* (Hs00995519\_g1), *Sox2* (NM\_011443, System Biosciences).

**Immunofluorescence of histones in MEFs and HSCs.** Cells were cultured for 48 h (for MEFs) or directly cytospinned on glass slides after purification (for HSCs). Cells were fixed in PFA 2% for 10 min and washed in Triton-PBS 0.1%. We performed a blocking incubation in PBS with donkey serum 10% for 1 h at room temperature (16–22 °C) before incubation with primary antibody at 4 °C overnight. Incubation with secondary antibody and DAPI was set for 45 min at room temperature. Alternatively, antibody staining was performed with Zenon kit (Invitrogen) following manufacturer's directions. Samples were then mounted and imaged.

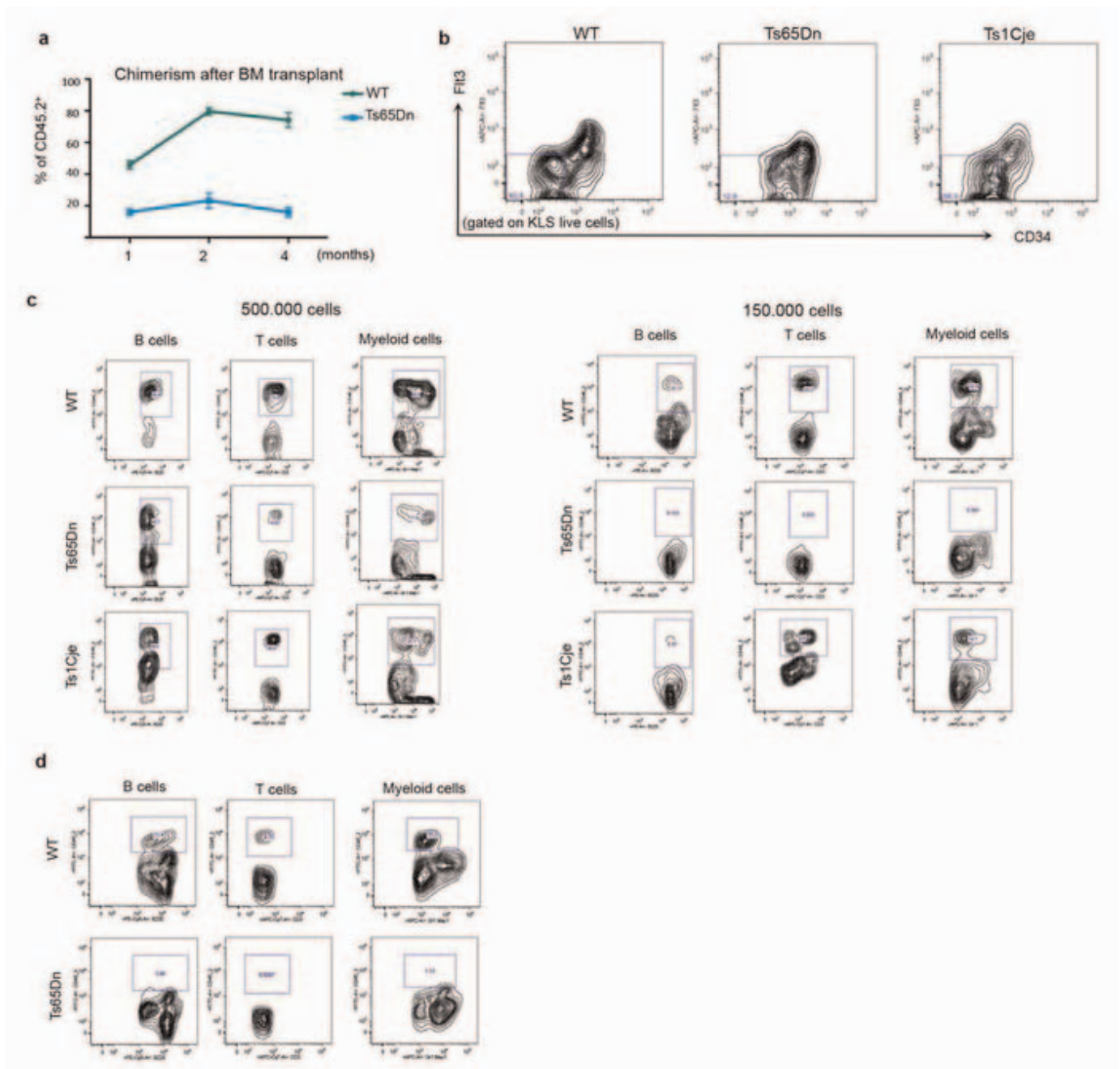
Imaging was performed using a Zeiss Observer Z1 fluorescent microscope (Zeiss) equipped with a Hamamatsu Orca-ER camera or a Zeiss confocal system LSM710 (Zeiss). Data acquisition and foci measurements were performed using Improvision Velocity software (Perkin Elmer).

The primary antibody used in this study was anti ubiquityl-histone H2A D27C4 (Cell Signaling). Secondary antibodies were Alexa Fluor 488 and Alexa Fluor 594 (Invitrogen).

**Human bone marrow samples.** Down's syndrome bone marrow samples were retrieved from sternal biopsies obtained from children undergoing corrective heart surgery at Stanford University. The families were consented according to IRB approved protocols. Control bone marrow samples were obtained from AllCells, LLC. All samples were processed and stained as described previously<sup>39</sup>.

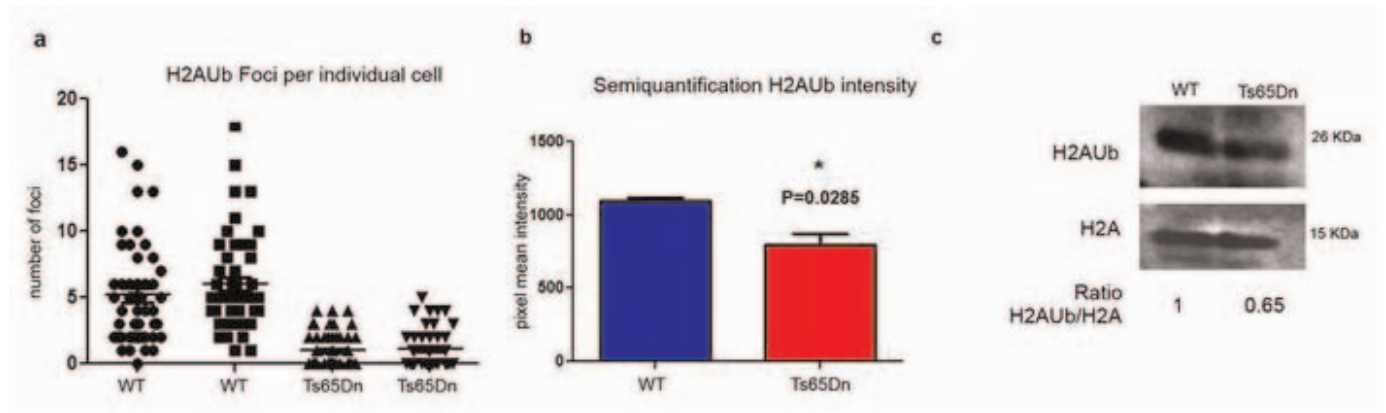
- Reinholdt, L. G. *et al.* Molecular characterization of the translocation breakpoints in the Down syndrome mouse model Ts65Dn. *Mamm. Genome* **22**, 685–691 (2011).
- Hu, Y. & Smyth, G. K. ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. *J. Immunol. Methods* **347**, 70–78 (2009).
- Dalerba, P. *et al.* Single-cell dissection of transcriptional heterogeneity in human colon tumors. *Nature Biotechnol.* **29**, 1120–1127 (2011).
- Zeng, Y. A. & Nusse, R. Wnt proteins are self-renewal factors for mammary stem cells and promote their long-term expansion in culture. *Cell Stem Cell* **6**, 568–577 (2010).
- Negishi, M. *et al.* A novel zinc finger protein Zfp277 mediates transcriptional repression of the *Ink4a/arf* locus through polycomb repressive complex 1. *PLoS ONE* **5**, e12373 (2010).
- Ventura, A. *et al.* Cre-lox-regulated conditional RNA interference from transgenes. *Proc. Natl Acad. Sci. USA* **101**, 10380–10385 (2004).
- Shimono, Y. *et al.* Downregulation of miRNA-200c links breast cancer stem cells with normal stem cells. *Cell* **138**, 592–603 (2009).
- Tiscornia, G., Singer, O. & Verma, I. M. Production and purification of lentiviral vectors. *Nature Protocols* **1**, 241–245 (2006).
- Majeti, R., Park, C. Y. & Weissman, I. L. Identification of a hierarchy of multipotent hematopoietic progenitors in human cord blood. *Cell Stem Cell* **1**, 635–645 (2007).





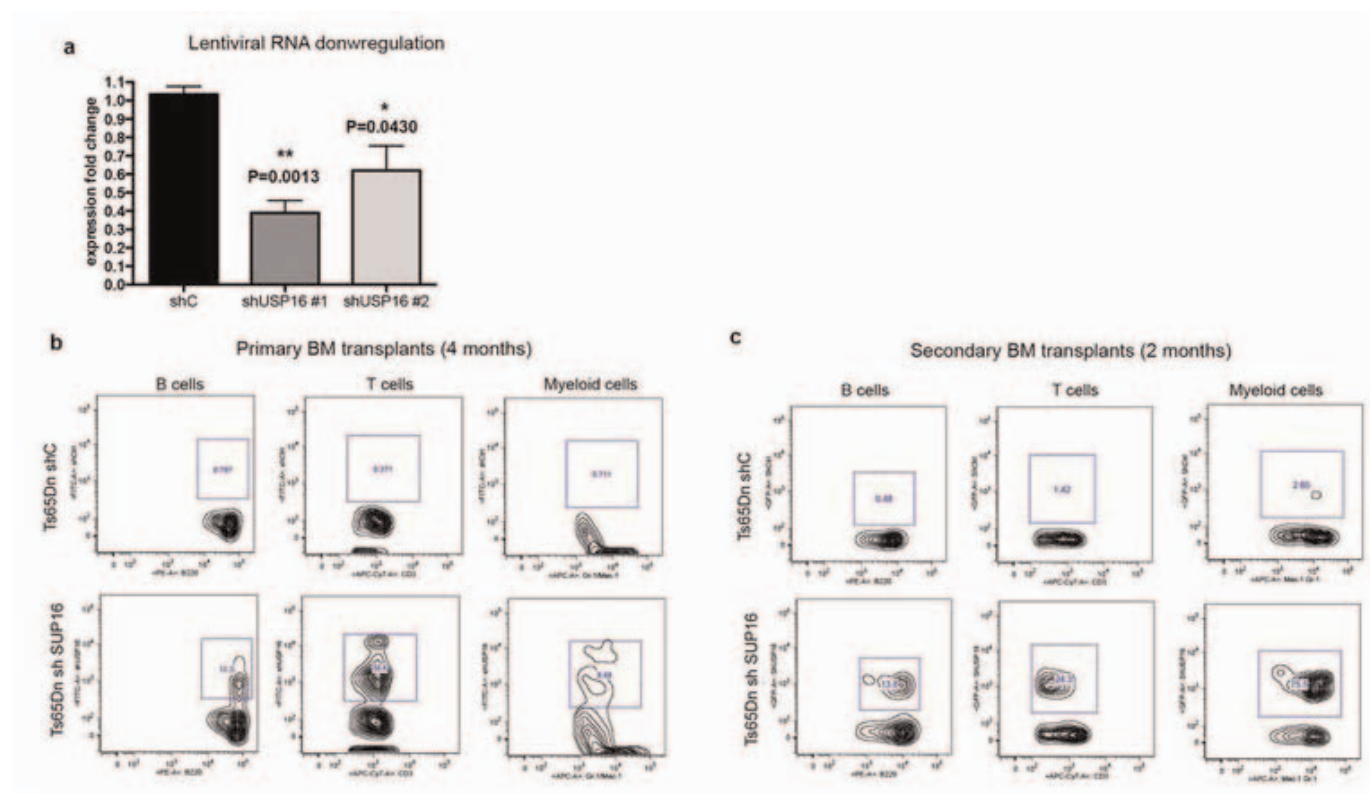
**Extended Data Figure 1 | Ts65Dn, but not Ts1Cje, bone marrow show a decreased number of functional HSCs.** **a**, Impaired engraftment ability of Ts65Dn bone marrow (BM) cells in transplantation experiments. The percentage of donor cells (CD45.2<sup>+</sup>) was evaluated at the indicated time points (five mice per group, repeated twice). **b**, Representative fluorescence-activated cell sorting (FACS) plots of KLS cells show a reduced number of CD34<sup>+</sup> Flt3<sup>+</sup> KLS cells in Ts65Dn marrows (four mice of each genotype were examined).

**c**, Peripheral blood analyses four months after bone marrow transplant revealed multi-lineage engraftment in Ts65Dn mice with only 500,000 donor cells; lower doses of Ts65Dn donor bone marrow failed to reconstitute haematopoietic lineages. **d**, Multi-lineage analyses of peripheral blood three months after bone marrow transplantation showed that Ts65Dn bone marrow cells failed to reconstitute secondary recipients. Representative FACS plots are shown.



**Extended Data Figure 2 | HSCs in Ts65Dn mice have lower levels of H2A ubiquitination.** **a**, Immunofluorescence for H2A ubiquitination on Lys 119 shows a decrease in the number of positive foci in MEFs derived from Ts65Dn embryos compared to controls. Each dot represents a different cell and each column a different mouse. One hundred cells per group were analysed and the

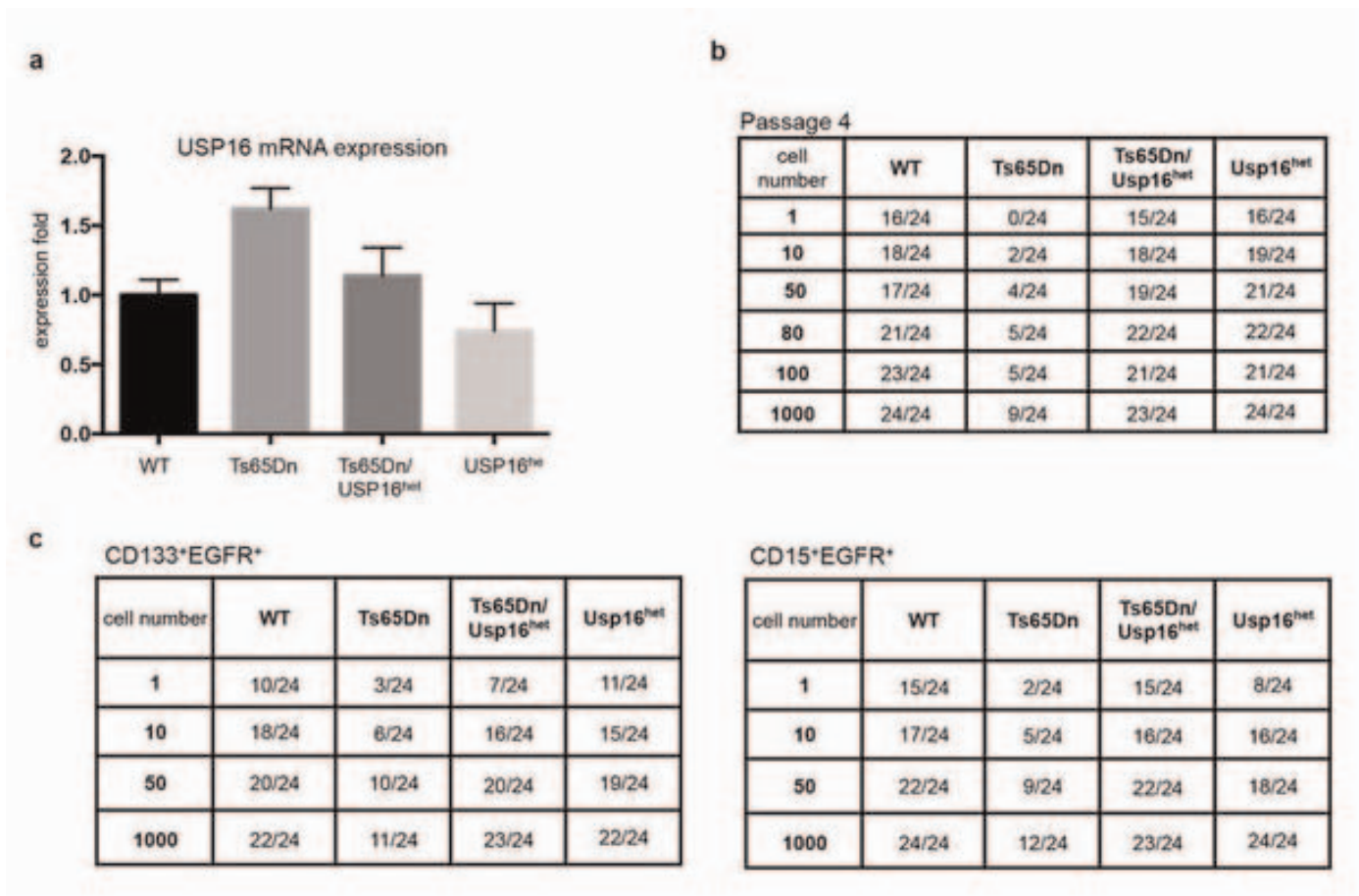
experiment was repeated twice. **b**, H2AK119<sup>+</sup> staining is decreased in Ts65Dn compared to control MEFs. **c**, Western blot analyses of chromatin extracts from MEFs. H2AK119<sup>+</sup> levels are decreased in Ts65Dn (quantification performed using ImageJ software). H2A western blotting verifies equal loading of extracts.



**Extended Data Figure 3 | Downregulation of Usp16 improves engraftment of Ts65Dn KLS cells in primary and secondary transplants. a, *Usp16* mRNA quantification after infection of KLS cells with the indicated lentivirus. b, Peripheral blood analyses revealed multi-lineage engraftment from Ts65 KLS**

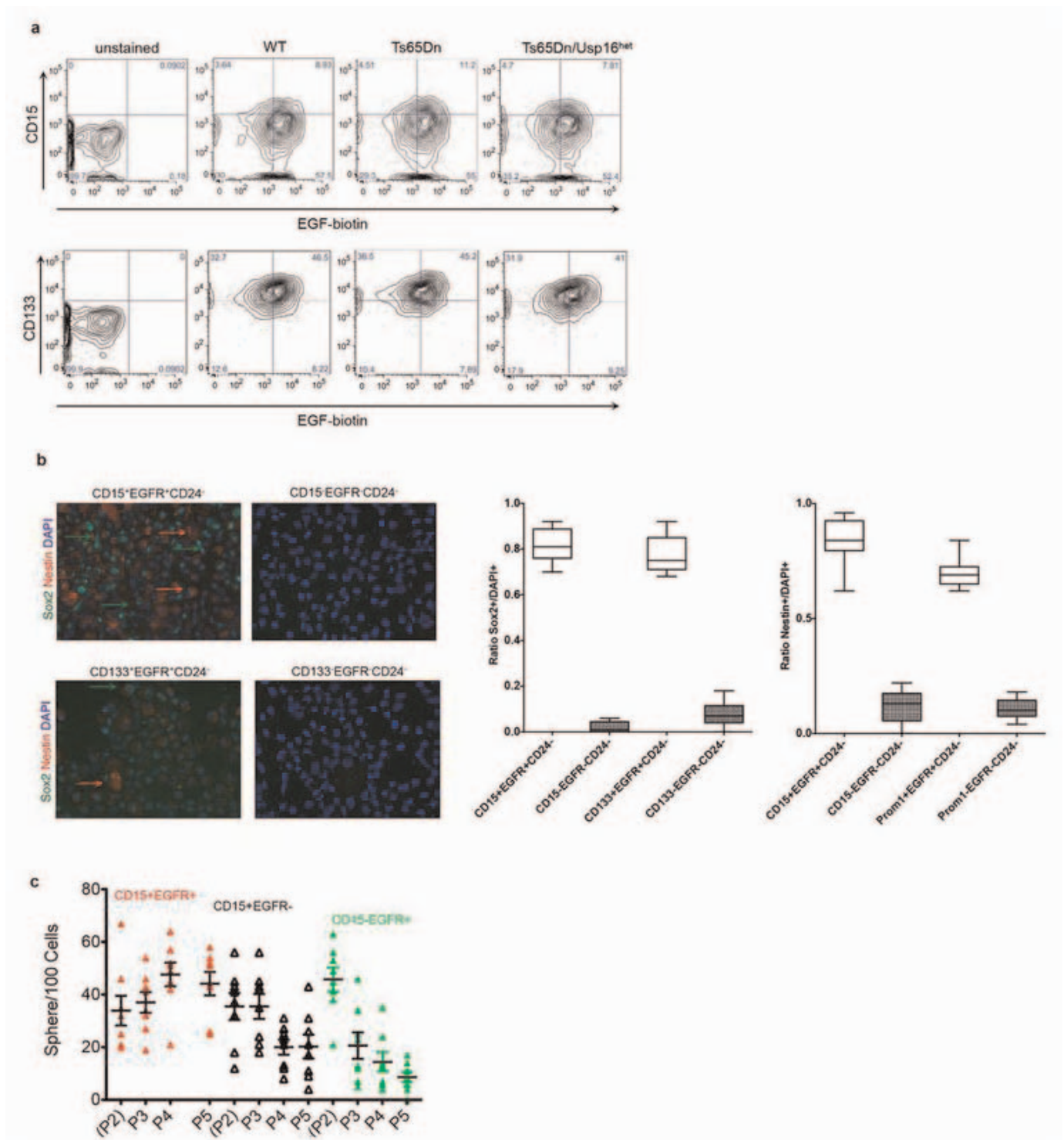
bone marrow cells infected with a *shUsp16* hairpin. Representative FACS plots are shown. c, Two months after transplantation in secondary recipients, *shC* Ts65Dn bone marrow cells fail to engraft, whereas *shUsp16* Ts65Dn cells show multi-lineage reconstitution. Representative FACS plots are shown.





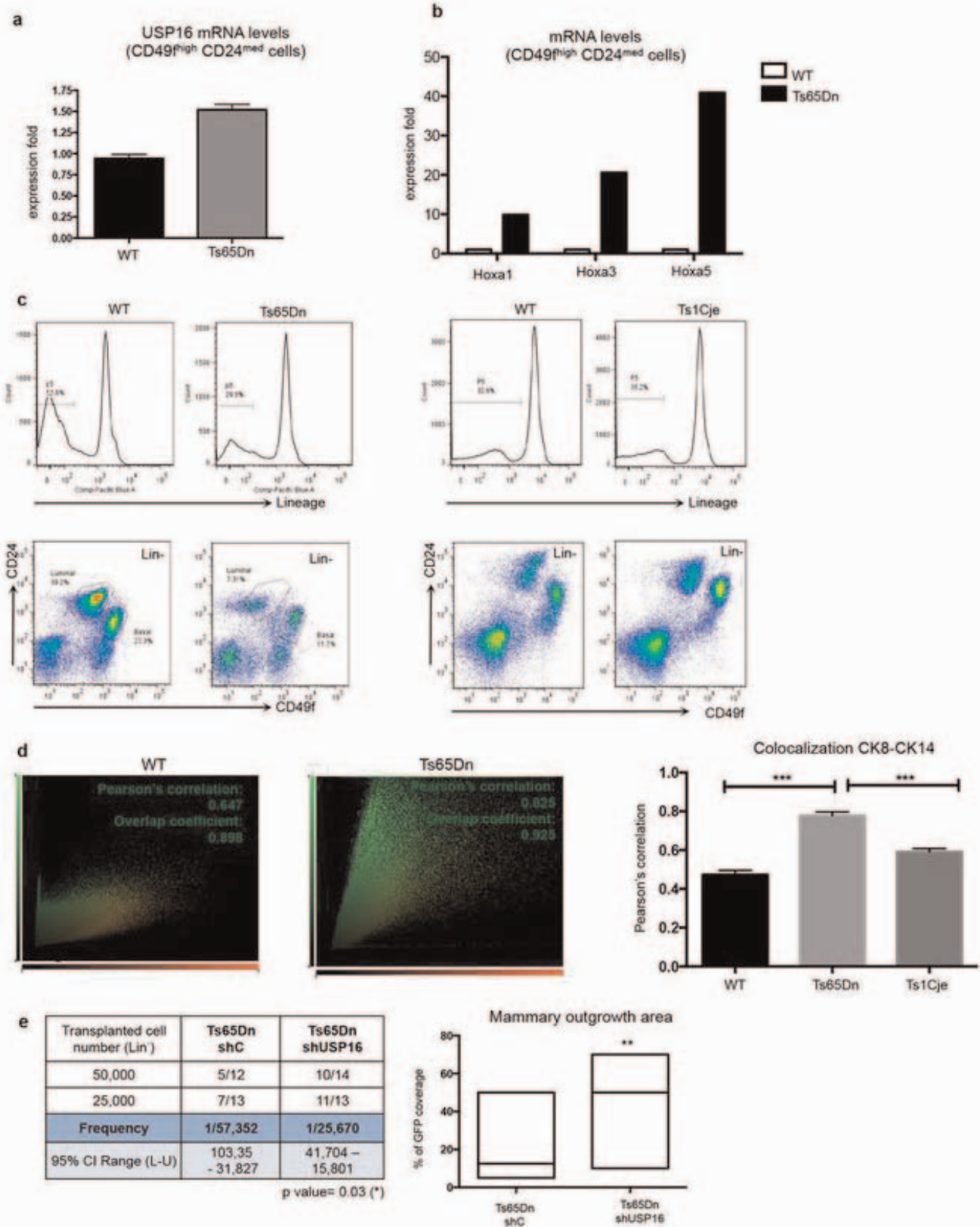
**Extended Data Figure 4 | Analyses of Nsp-IC frequency in neurospheres cultures.** **a**, *Usp16* mRNA quantification in murine neurospheres cultures (P4). **b**, **c**, Raw data used for ELDA analyses of Nsp-ICs derived from Lin<sup>−</sup> SVZ cells

or for the indicated sorted population. For each cell dilution, 24 replicates were tested. The table indicates the number of positive wells in each condition.



**Extended Data Figure 5 |  $CD15^+ Egfr^+$  and  $Prom1^+ Egfr^+$  populations are enriched for neuronal progenitors in mice.** **a**, Representative FACS plots are shown for viable  $Lin^-$  cells derived from SVZ preparations. Double-positive cells were sorted and used for testing neurosphere-formation potential. **b**, Representative pictures of immunofluorescence staining for Sox2 and nestin on the indicated sorted populations. The arrows indicate cells scored positive for Sox2 (green) or nestin (red). For this analysis, the indicated  $Lin^-$  cell

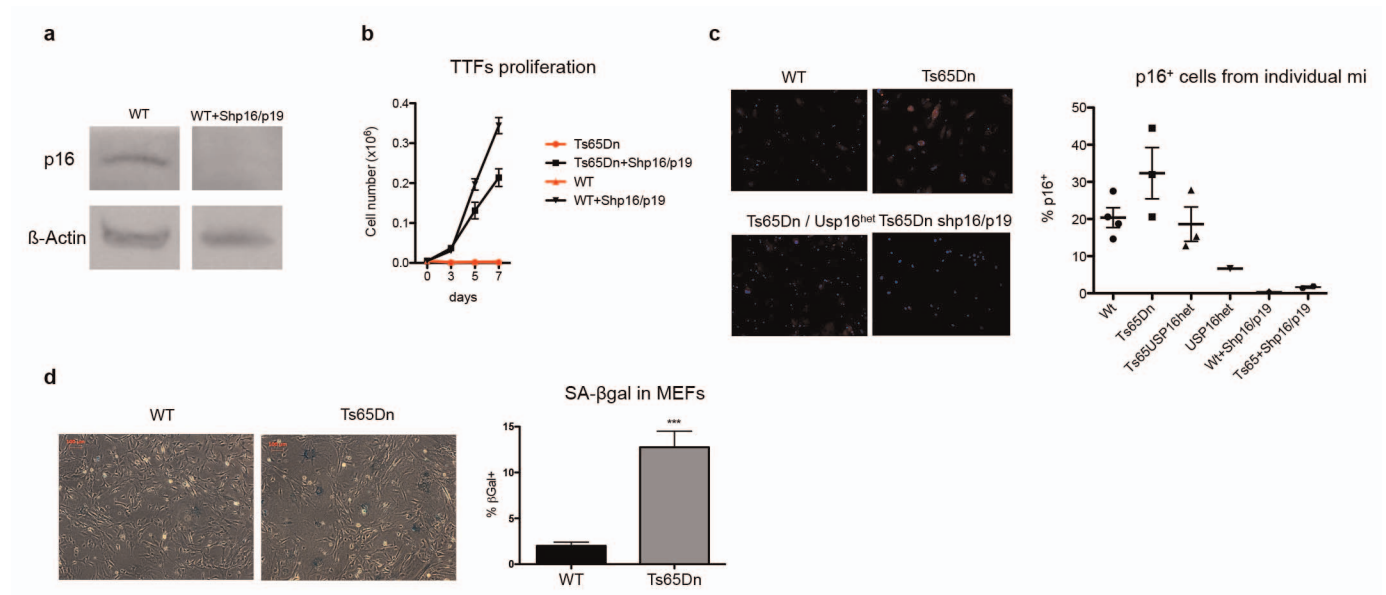
populations were FACS sorted and collected by cytopsin. On the right, twelve fields were randomly selected for analyses from four wild-type mice from different litters. The percentage of positive cells is given by the ratio of cells positive for Sox2 or nestin among the DAPI (4',6-diamidino-2-phenylindole) $^+$  cells. **c**, Neurosphere expansion *in vitro* during passaging by different sorted populations derived from mouse SVZ.  $CD15^+ Egfr^+$  cells are able to expand upon passaging.





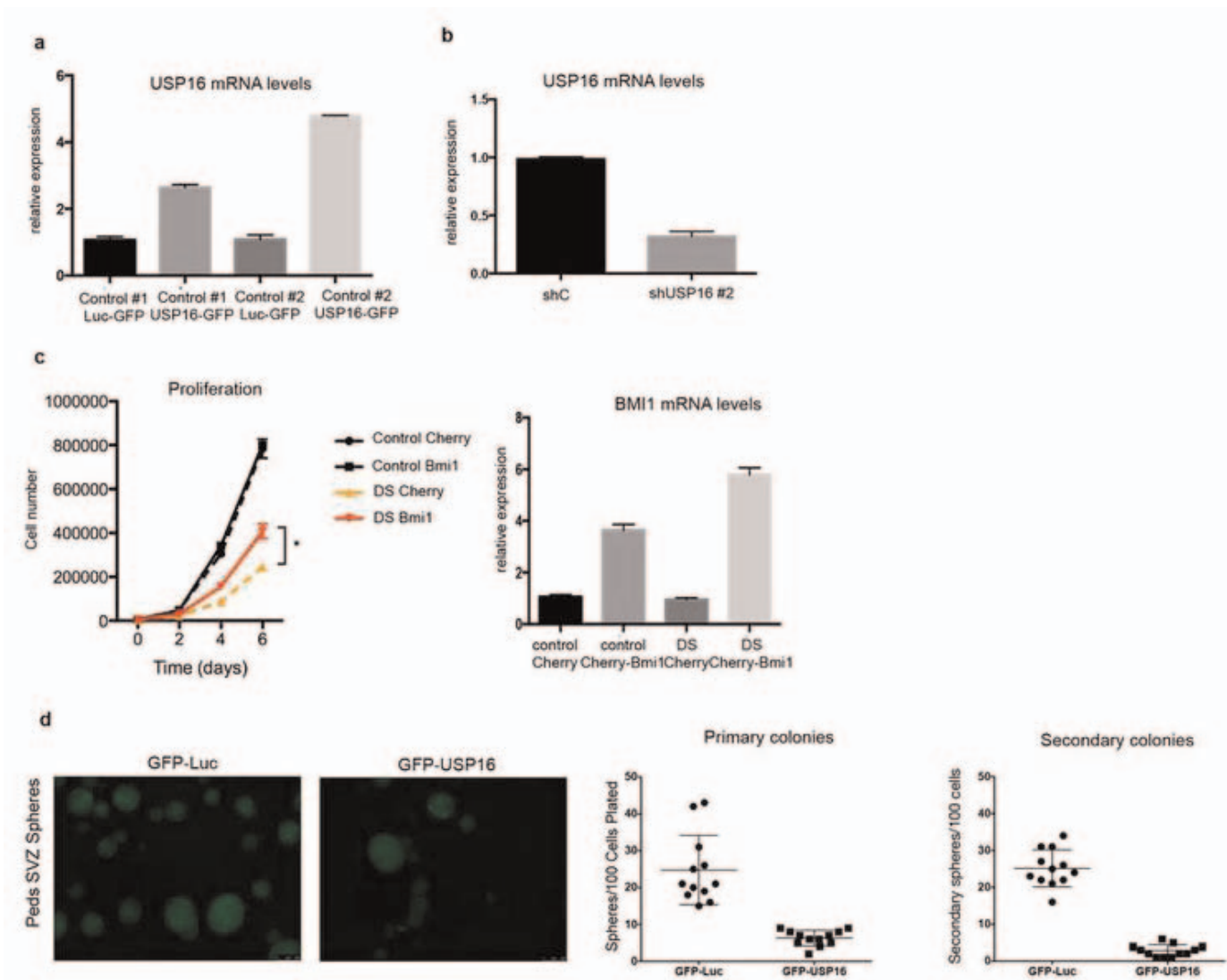
**Extended Data Figure 6 | Defects in mammary glands in Down's syndrome mice models.** **a, b**, mRNA quantification of *Usp16* and different *Hox* genes in CD49<sup>high</sup> CD24<sup>med</sup> mammary cells. *Hox1*, *Hox3* and *Hox5* are expressed at higher levels in Ts65Dn cells. **c**, Representative FACS plot of mammary cells gated on live cells (first row) or live Lin<sup>−</sup> cells. We observed a perturbation in the overall FACS profile with reduction of basal and luminal cells (indicated gates) in Ts65Dn mice but not in Ts1Cje mice. These experiments were repeated five times for each group. **d**, Quantification of overlap between staining for the basal cytokeratin CK14 (red) and the luminal cytokeratin CK8

(green). Pearson's correlation analyses (luminosity software) showed a marked increase in cells that co-stain for both cytokeratins in Ts65Dn mammary epithelium. Each experiment was repeated with three mice per group. **e**, Downregulation of *Usp16* by shRNA lentiviral infection partially rescues the *in vivo* defects shown by Ts65Dn mammary cells ( $P = 0.03$ ). Three independent transplantation experiments were performed. Right, the percentage of fat pad filled by GFP outgrowths is significantly higher upon downregulation of *Usp16* ( $P = 0.007$ ).



**Extended Data Figure 7 | Senescence in Ts65Dn fibroblasts is affected by levels of Usp16 and Cdkn2a.** **a**, Western blot analyses verifies knockdown of p16. β-actin was used as a loading control. **b**, Proliferation of Ts65Dn TTFs increased upon infection with a hairpin targeting *Cdkn2a*. Control TTFs proliferate more upon downregulation of both *p16<sup>Ink4a</sup>* and *p19<sup>Arf</sup>*. The experiment was repeated three times with similar results. **c**, Representative

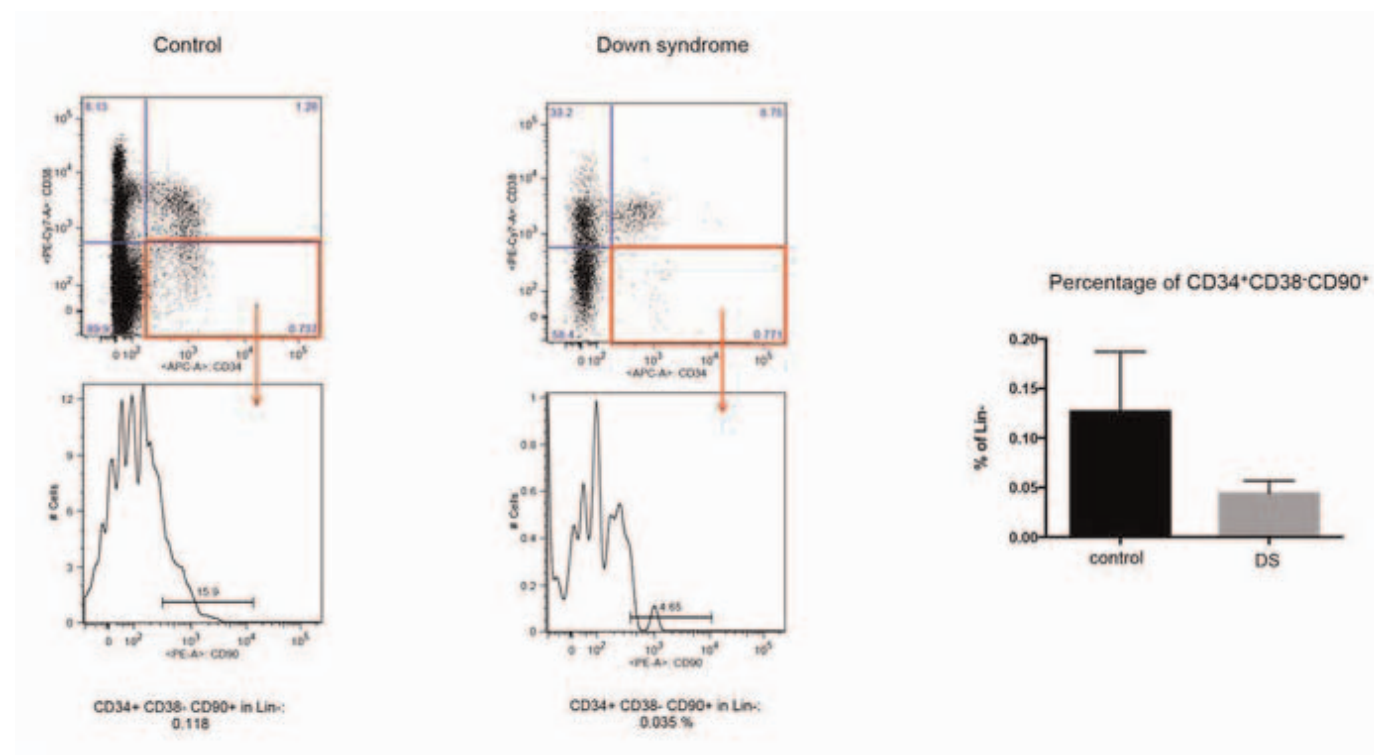
pictures of p16<sup>Ink4a</sup> immunostaining (left) and quantification of the percentage of positive cells (right). Each dot represents a TTF culture derived from a different mouse. The hairpin effectively ablates p16<sup>Ink4a</sup> expression. **d**, SA-βgal staining in control and Ts65Dn MEFs at P4. Representative pictures are shown on the left. The percentages of positive cells are shown on the right. Experiments were replicated with three different MEF lines per genotype.



**Extended Data Figure 8 | Human fibroblast cultures.** **a, b**, mRNA quantification of *USP16* upon lentiviral overexpression or downregulation. **c**, *BMI1* overexpression significantly increases the proliferation of fibroblasts derived from a Down's syndrome carrier. The effect on control fibroblasts is not significant. On the right, the levels of expression of *BMI1* mRNA were

quantified by real-time PCR. The experiment was repeated twice with similar results. **d**, Overexpression of *USP16* reduces the formation of neurospheres derived from human paediatric SVZ cells. On the right, the quantification of the number of spheres obtained in the first and in the second passage. All the experiments were replicated twice.





**Extended Data Figure 9 | HSCs in human bone marrow derived from Down's syndrome patients.** Left, representative plot of Lin<sup>-</sup> cells derived from human bone marrow. Right, a plot of the percentage of CD34<sup>+</sup>CD38<sup>-</sup>CD90<sup>+</sup> observed between the Lin<sup>-</sup> cells (Down's syndrome: two samples; healthy donors: three samples). Down's syndrome bone marrow samples were retrieved from sternal biopsies obtained from children undergoing corrective heart surgery at Stanford University. The families were consented according to IRB approved protocols. Control bone marrow samples were obtained from AllCells, LLC. Although we cannot unequivocally state that the HSCs are reduced in Down's syndrome patients, in these children they are not increased as they might be in the liver of some human Down's syndrome fetuses. Down's syndrome is associated with increased rates of childhood leukaemia and decreased rates of some adult solid tumours. We can only speculate as to why

Down's syndrome predisposes individuals to one type of malignancy and protects from another, and whether USP16 has a role in either observation. Other syndromes causing bone marrow failure, such as Fanconi anaemia, predispose to leukaemia. Lymphoid leukaemias in Down's syndrome patients frequently involve mutation of CDKN2A. Because CDKN2A appears to have a role in the proliferation defects caused by trisomy of *USP16*, mutations of *CDKN2A* could give Down's syndrome HSCs a strong selection advantage. Differences between fetal liver and adult bone marrow HSCs could provide another explanation for an increased incidence of leukaemia in Down's syndrome. An expansion of Down's syndrome fetal liver HSCs but not of adult bone marrow HSCs could also explain why Down's syndrome leukaemia appears to originate more often in fetal blood cells.

# Structural insight into the biogenesis of $\beta$ -barrel membrane proteins

Nicholas Noinaj<sup>1</sup>, Adam J. Kuszak<sup>1</sup>, James C. Gumbart<sup>2</sup>, Petra Lukacik<sup>3</sup>, Hoshing Chang<sup>1</sup>, Nicole C. Easley<sup>1</sup>, Trevor Lithgow<sup>4</sup> & Susan K. Buchanan<sup>1</sup>

**$\beta$ -barrel membrane proteins are essential for nutrient import, signalling, motility and survival. In Gram-negative bacteria, the  $\beta$ -barrel assembly machinery (BAM) complex is responsible for the biogenesis of  $\beta$ -barrel membrane proteins, with homologous complexes found in mitochondria and chloroplasts. Here we describe the structure of BamA, the central and essential component of the BAM complex, from two species of bacteria: *Neisseria gonorrhoeae* and *Haemophilus ducreyi*. BamA consists of a large periplasmic domain attached to a 16-strand transmembrane  $\beta$ -barrel domain. Three structural features shed light on the mechanism by which BamA catalyses  $\beta$ -barrel assembly. First, the interior cavity is accessible in one BamA structure and conformationally closed in the other. Second, an exterior rim of the  $\beta$ -barrel has a distinctly narrowed hydrophobic surface, locally destabilizing the outer membrane. And third, the  $\beta$ -barrel can undergo lateral opening, suggesting a route from the interior cavity in BamA into the outer membrane.**

Membrane proteins serve numerous essential functions and are important therapeutic targets given their surface exposure and crucial roles in modulating cellular processes. Although the mechanism for membrane integration is well established for  $\alpha$ -helical membrane proteins<sup>1–4</sup>, the mechanism for  $\beta$ -barrel membrane proteins is unknown.  $\beta$ -barrel membrane proteins are only found in the outer membranes of Gram-negative bacteria and mitochondria and chloroplasts, eukaryotic organelles that evolved from bacteria<sup>5–10</sup>. The machineries catalysing folding and insertion of  $\beta$ -barrel proteins have been identified and are conserved across species<sup>6,9,10</sup>.

In Gram-negative bacteria, outer membrane proteins (OMPs) are synthesized in the cytoplasm and transported across the inner membrane into the periplasm by the Sec translocon<sup>11</sup>. Molecular chaperones then escort nascent OMPs to the inner surface of the outer membrane, where they are recognized by the BAM complex, which consists of the central and essential component known as BamA (an OMP itself) and the accessory proteins BamB, BamC, BamD and/or BamE, all of which reside in the periplasm and are attached to the membrane via a lipid anchor<sup>6,12–16</sup> (Supplementary Fig. 1). The periplasmic domain of BamA consists of five polypeptide translocation-associated (POTRA) domains that extend from the barrel. Current understanding suggests that the four lipoproteins assemble onto the POTRA scaffold to create a BAM complex consisting of one copy of each protein<sup>14,17</sup>. Structures have been determined for BamB<sup>18–20</sup>, BamC<sup>20,21</sup>, BamD<sup>20–23</sup>, BamE<sup>20,21,24</sup> and the periplasmic (POTRA) domain of BamA<sup>25–29</sup>. However, an understanding of how these proteins coordinate recognition, folding and membrane insertion of nascent OMPs has been hampered by a lack of structural knowledge of the membrane domain of BamA.

To understand biogenesis of OMPs better, we solved crystal structures of BamA from two bacterial species: *N. gonorrhoeae* and *H. ducreyi*. Two distinct POTRA domain conformations are observed relative to the  $\beta$ -barrel, with the barrel pore either fully accessible from the periplasm or blocked by a POTRA domain. The structures also reveal that BamA has a reduced external hydrophobic surface on one side that could produce local distortions in the outer membrane. Intriguingly, whereas the *H. ducreyi* BamA is in a predictable conformation with the

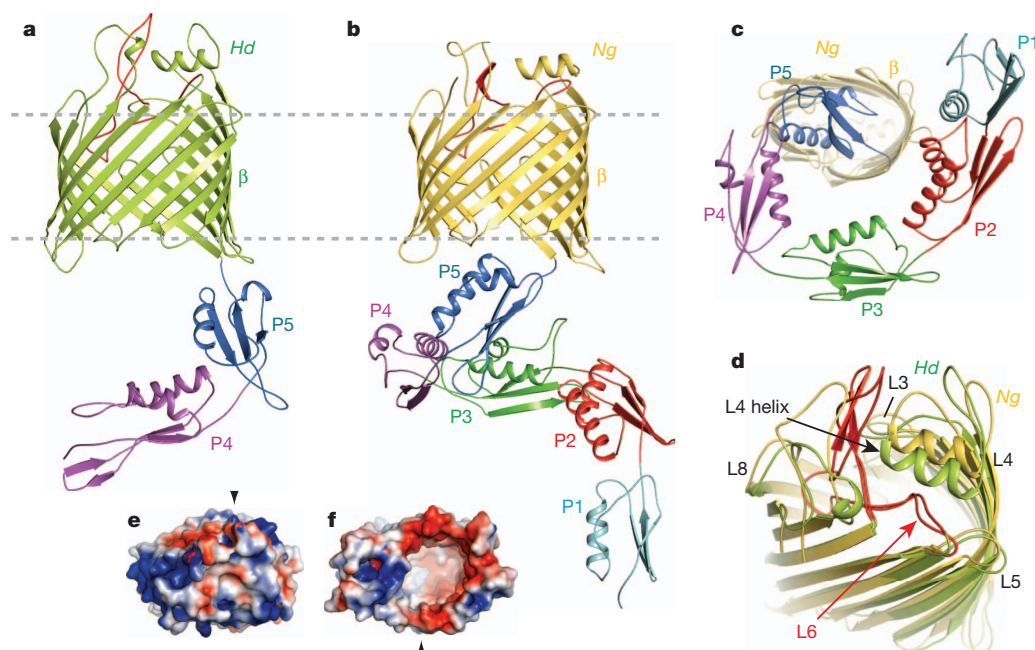
first and last  $\beta$ -strands stably zipped with multiple hydrogen bonds, the last  $\beta$ -strand of the *N. gonorrhoeae* BamA structure is bound to the first by only two hydrogen bonds, with most of the  $\beta$ -strand twisted into the barrel pore. In this conformation, a large access portal provides a direct connection from the periplasm to the lipophilic interior of the outer membrane.

## Two structures of BamA

We determined structures for an amino-terminally truncated BamA construct from *H. ducreyi* lacking the first three POTRA domains (*HdBamA* $\Delta$ 3), and a full-length BamA construct from *N. gonorrhoeae* (*NgBamA*) (Fig. 1a, b, Supplementary Table 1 and Supplementary Figs 2–4). Both structures include the large carboxy-terminal  $\beta$ -barrel membrane domain (Fig. 1a, b). Despite sequence divergence, each of the POTRA domains retains the conserved  $\beta$ - $\alpha$ - $\beta$  fold (Supplementary Figs 5 and 6). In the *HdBamA* $\Delta$ 3 structure, the POTRA domains extend away from the barrel, allowing full access to the barrel pore from the periplasm. By contrast, the POTRA domains of *NgBamA* are located in close proximity to the barrel pore, such that POTRA domain 5 (POTRA5) occludes pore access (Fig. 1b, c).

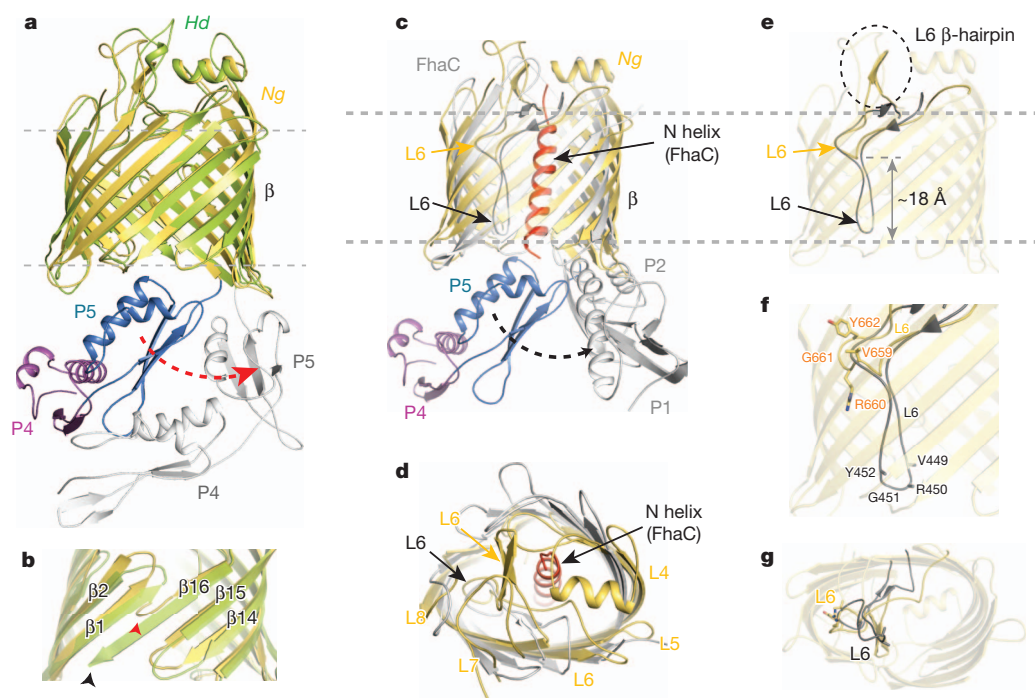
The barrel domains of *HdBamA* $\Delta$ 3 and *NgBamA* each contain 16 antiparallel  $\beta$ -strands, with the first and last strands associating to close the barrel (Fig. 1a, b). Notably, the interior of the barrel is almost completely empty, forming a volume of  $\sim 13,000 \text{ \AA}^3$  (Supplementary Figs 8 and 9). The extracellular loops form a dome over the top of the  $\beta$ -barrel domain, isolating the inside of the barrel from the extracellular space (Fig. 1d). Extracellular loops (eL) eL4, eL6 and eL7 contribute substantially to the dome, with minor contributions from eL3 and eL8. Extracellular loop eL4 contains a surface-exposed  $\alpha$ -helix that sits nearly parallel to the membrane. Although the sequences for this helix vary considerably among BamA homologues, it is structurally conserved (Fig. 1d and Supplementary Fig. 10). The electrostatics of the surface-exposed regions in both structures contain a mix of positive and negative patches, with a strongly electropositive surface along eL3 and eL6 (Fig. 1e, f) and electropositive charge at the membrane interfaces. The electrostatics of the barrel interior are similar in the two structures,

<sup>1</sup>National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA. <sup>2</sup>School of Physics, Georgia Institute of Technology, Atlanta, Georgia 30332, USA. <sup>3</sup>Diamond Light Source Ltd, Oxfordshire OX11 0DE, UK. <sup>4</sup>Department of Biochemistry and Molecular Biology, Monash University, Victoria 3800, Australia.



**Figure 1 | The structure of BamA from the BAM complex.** **a**, The *HdBamA*Δ3 crystal structure in cartoon representation showing the β-barrel (green) and POTRA domains 4 and 5 (purple and blue, respectively). **b**, The *NgBamA* crystal structure showing the β-barrel (gold) and POTRA domains 1–5 (cyan, red, green, purple and blue, respectively). **c**, A periplasmic (bottom) view of the *NgBamA* crystal structure. **d**, An alignment of the *HdBamA*Δ3

(green) and *NgBamA* (gold) crystal structures highlighting the structural conservation of the extracellular loops and secondary structural elements in loops (L) 4 and 6. **e**, **f**, Electrostatic surface representation of *HdBamA*Δ3 viewed from the extracellular face (**e**) and the periplasmic face (**f**). Black arrowheads indicate the locations of strand β16.



**Figure 2 | *HdBamA* and *NgBamA* crystal structures reveal conformational changes.** **a**, Alignment of *HdBamA*Δ3 (green) and *NgBamA* (gold) showing open and closed conformations for the POTRA domains, which may serve as a gating mechanism for regulating substrate access to the inside of the β-barrel. **b**, Compared to *HdBamA*Δ3 (green), β-strand 16 is disordered and tucked inside the β-barrel of *NgBamA* (gold). Arrowheads indicate the location of the C-terminal strand in *HdBamA* (black) and *NgBamA* (red). **c**, **d**, Membrane (**c**) and extracellular (**d**) view of an alignment of *NgBamA* and FhaC (grey, Protein Data Bank (PDB) code 2QDZ) illustrates conformational differences in the β-barrel and POTRA domains. In FhaC, the N-terminal α-helix (red) and

loop 6 occlude the β-barrel preventing free diffusion across the outer membrane; however, in BamA this is accomplished by the extracellular loops that fold over the top of the barrel. **e**, **f**, Extracellular loop 6 (eL6) assumes different conformations between the two structures: membrane view (**e**), zoomed membrane view (**f**), and periplasmic view (**g**). Unlike FhaC, *NgBamA* eL6 contains a β-hairpin (dashed circle), and the VRGF/Y motif is located ~18 Å from the periplasmic boundary. Unzipping of this β-hairpin may allow for an extended conformation similar to what is observed for eL6 of FhaC.



both displaying strongly electronegative surfaces (Fig. 1f and Supplementary Figs 8 and 9).

## Two conformations of BamA

Although the two BamA crystal structures show conserved folds for both the POTRA domains and the  $\beta$ -barrel, there are notable differences between the two species. First is the conformation of the POTRA domains relative to the  $\beta$ -barrel (Fig. 2a). In the NgBamA structure, POTRA5 sits in proximity to the  $\beta$ -barrel and interacts with periplasmic loops (pL) pL3, pL4, pL5 and pL7 to stabilize this closed conformation (Supplementary Fig. 7). Conversely, the POTRA domains of the HdBamA $\Delta$ 3 structure have undergone a  $\sim 70^\circ$  outward swing such that POTRA5 does not interact with any of the  $\beta$ -barrel periplasmic loops. This could reflect a gating mechanism for regulating access to the interior of the  $\beta$ -barrel.

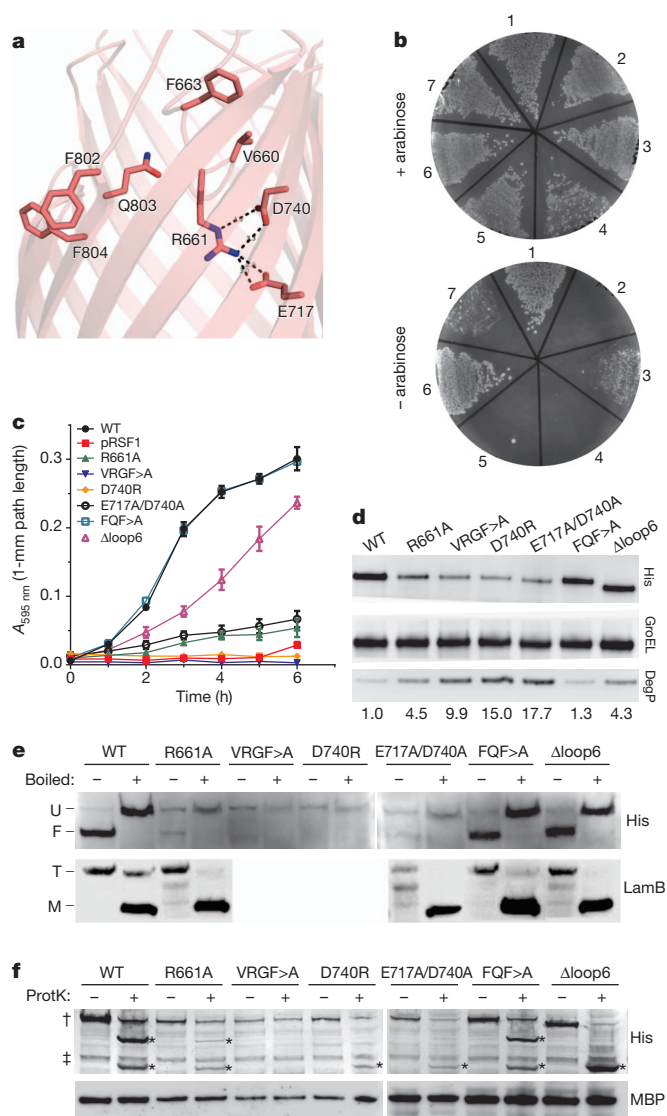
A second difference is found at the interface of strands 1 and 16 of the barrel. In the HdBamA $\Delta$ 3 structure, strands 1 and 16 associate to close the  $\beta$ -barrel with eight hydrogen bonds (Fig. 2b and Supplementary Fig. 7). However, in the NgBamA structure the C-terminal  $\beta$ -strand is twisted and tucked inside the barrel, interacting with strand 1 via only two hydrogen bonds at the extracellular face of the barrel. The NgBamA structure provides the first example, to our knowledge, of destabilization of the C-terminal  $\beta$ -strand, and this conformation would allow access from the interior cavity of BamA to the lipid phase of the outer membrane at the interface of strands 1 and 16.

## Comparison of BamA to FhaC

Until now, the only source of structural information for the membrane domain of any member of the Omp85 family of proteins has come from FhaC, which serves as a dedicated toxin translocation pore in the outer membrane of some bacteria<sup>30,31</sup>. BamA and FhaC share less than 13% sequence identity and function in very different processes, so it is not surprising that the structures differ greatly. First, the root mean squared deviation (r.m.s.d.) values for the 16-stranded  $\beta$ -barrels are larger than 10 Å, owing to differing shear numbers ( $S = 20$  for FhaC and  $S = 22$  for HdBamA and NgBamA) and overall barrel shapes (Fig. 2c). Second, although the extracellular loops of FhaC are in open conformations, as befits a toxin translocation pore, the extracellular loops of HdBamA $\Delta$ 3 and NgBamA form a closed dome that prevents extracellular access and would limit free diffusion of solutes across the outer membrane (Fig. 2d). Third, the conformation of eL6 differs considerably for FhaC and the two BamA structures. This loop has gained much attention owing to its large size and the location of the conserved VRGF/Y (Val-Arg-Gly-Phe/Tyr) motif, with the suggestion that eL6 may extend through the barrel pore to assume at least two conformations, one near the periplasm and another closer to the surface<sup>30,32</sup>. In both BamA crystal structures, eL6 partially inserts into the barrel pore such that the VRGF/Y motif interacts with  $\beta$ -strands 14–16 about 18 Å away from the periplasmic side of the outer membrane (Fig. 2e and Supplementary Fig. 7). This interaction is mediated by a conserved arginine residue in the barrel pore (Arg658 in HdBamA $\Delta$ 3 and Arg660 in NgBamA). The arginine is stabilized by interactions with conserved barrel residues Glu696 and Asp719 in HdBamA $\Delta$ 3 and Glu692 and Asp713 in NgBamA. Extracellular loop eL6 is further stabilized by interactions with a conserved FQF motif located in  $\beta$ -strand 16 (Supplementary Fig. 7). In FhaC, eL6 has a similar interaction with the  $\beta$ -barrel, but it extends through the pore, such that the VRGF/Y motif protrudes into the periplasm (Fig. 2f, g).

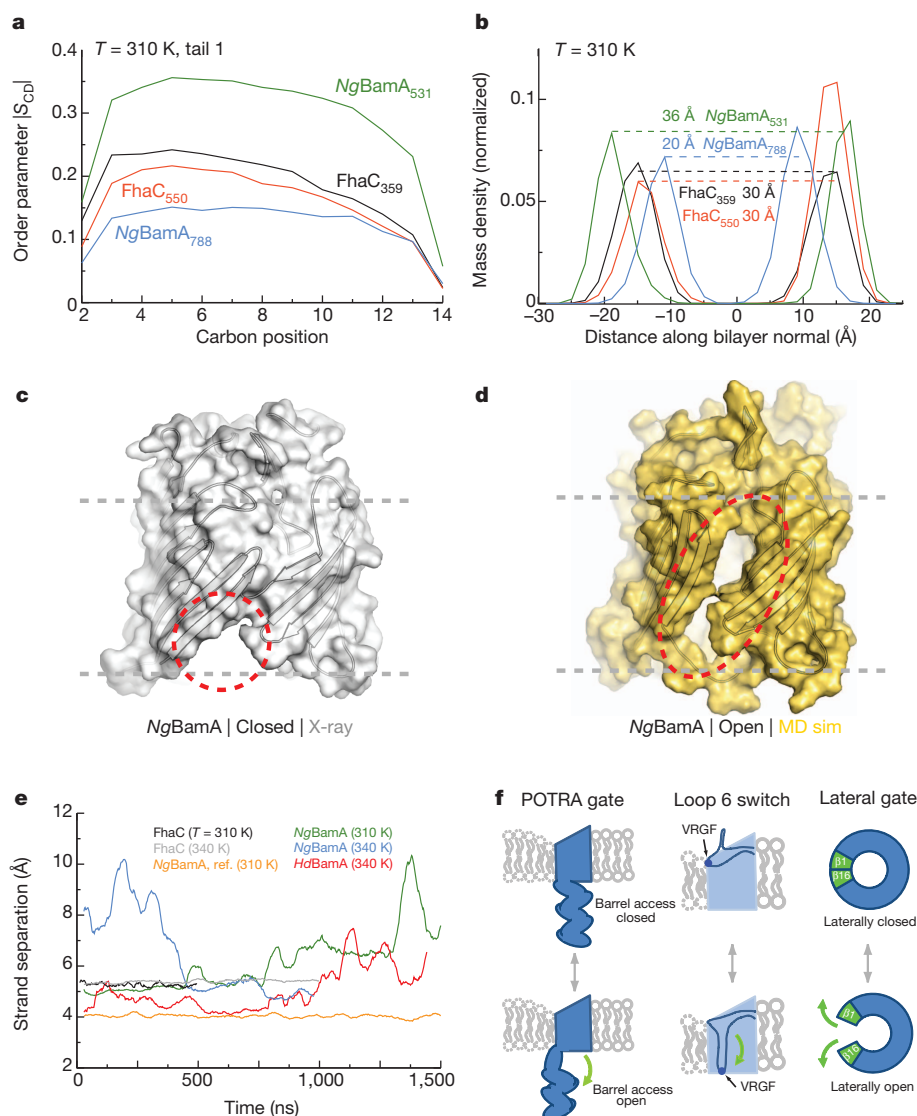
## Modelling and mutagenesis of EcBamA

Because most functional analyses of  $\beta$ -barrel assembly have been undertaken in *Escherichia coli*, we built a homology model of EcBamA (Supplementary Figs 10 and 11 and Supplementary Information) and mutated residues that are conserved across species or predicted by the model to be functionally relevant (Fig. 3a and Supplementary Table 2). Consistent with previous work highlighting the importance of the VRGF/Y



**Figure 3 | Mutational analysis of *E. coli* BamA.** **a**, A homology model of EcBamA showing the conserved VRGF/Y and FQF motifs and putative Arg661 ionic interactions (dashed lines). **b**, Colony growth assays of cells reliant on mutant EcBamA in the absence of arabinose: (1) wild-type EcBamA; (2) pRSF1 vector control; (3) Arg661Ala; (4) VRGF>A; (5) Asp740Arg; (6) Glu717Ala/Asp740Ala; and (7)  $\Delta$ eL6 (residues 676–700). **c**, Growth curves of colonies isolated from arabinose-containing plates and then transferred to medium in the absence of arabinose. Error bars denote  $\pm$  s.e.m. WT, wild type. **d**, Western blots of EcBamA mutant expression levels (anti-His) and DegP upregulation when grown in M63 minimal media in the absence of arabinose. GroEL served as a loading control; numbers represent DegP fold increases over wild type. **e**, Heat modifiability of EcBamA mutants expressed in LB containing arabinose, and of LamB after arabinose wash-out (at the 4 h time point as in **c**, in which no growth was observed for the vector control, VRGF>A, or Asp740Arg). All His-tagged EcBamA mutants except VRGF>A and Asp740Arg showed evidence of folding and LamB trimer formation. F, folded; M, monomer; T, trimer; U, unfolded. **f**, Anti-His reactive cleavage products (asterisks) after proteinase K (protK) proteolysis of whole cells indicated extracellular accessibility of all EcBamA mutants except VRGF>A. Dagger indicates full-length mutant BamA; double dagger indicates a nonspecific anti-His-horseradish-peroxidase reactive protein. Maltose binding protein (MBP) was used to control for outer membrane integrity. All data shown are representative of at least three independent experiments.

motif<sup>33</sup>, the Arg661Ala mutant exhibited reduced colony growth in LB-rich medium, and mutation of the entire VRGF/Y motif was lethal (Fig. 3b, c and Supplementary Table 2). Mutating Asp740 to arginine was also lethal, and the Glu717Ala/Asp740Ala double mutant demonstrated



**Figure 4 | BamA primes the membrane for OMP insertion.** Molecular dynamics simulations investigated the stability of the BamA  $\beta$ -barrel. **a**,  $S_{CD}$  values, a measure of lipid order, were decreased near  $NgBamA$   $\beta$ -strand 16 (centred at residue 788; blue) compared to the opposite side of the  $\beta$ -barrel (centred at residue 531; green). Minimal differences were observed for  $FhaC$  comparing analogous sites (black, red). **b**, Molecular dynamics analysis revealed that the  $\beta$ -barrel of  $NgBamA$  imparts a thinning of the membrane by 16  $\text{\AA}$  near strand  $\beta$ 16 (centred at residue 788) when compared to the opposite side of the barrel (centred at residue 531), whereas no difference was observed for  $FhaC$ . **c**, **d**, The molecular dynamics (MD) simulations also indicated that the interaction between  $\beta$ -strands 1 and 16 is severely destabilized, allowing the  $\beta$ -barrel to undergo a lateral opening (**c** and **d**, closed and open states,

minimal growth in shaking culture growth assays in LB media, confirming the proposed electrostatic interaction of the VRGF/Y motif with the inner barrel wall (Fig. 3c). Mutagenesis of potential interactions between the periplasmic loops and POTRA5 had no effect on cell viability, nor did mutation of the highly conserved FQF motif (Supplementary Table 2 and Supplementary Figs 10 and 12). Preventing a potential disulphide bond in eL6 had no effect in growth assays, but deletion of the non-conserved *E. coli* loop insertion (residues 676–700) resulted in reduced colony formation and a slower doubling time than the wild type (Fig. 3b, c).

Investigating these growth phenotypes, we found that the Arg661Ala, VRGF>A, Asp740Arg and Glu717Ala/Asp740Ala BamA mutants exhibited low expression levels compared to wild type (Fig. 3d). Furthermore, expression of these mutants and the  $\Delta$ loop6 deletion mutant

respectively). **e**, Quantification of the separation between  $\beta$ -strands 1 and 16 shows that both the  $NgBamA$  (green) and  $HdBamA$  (red) structures have the propensity to open, unlike  $FhaC$  (black and grey). As reference, no change was observed between  $\beta$ -strands 13 and 14 for  $NgBamA$  (orange). **f**, Summary of the putative conformational switches of BamA based on structural and computational analysis. The first is the conformational gate of the POTRA domains (membrane view), the second is the conformational switch of loop 6 from a resting state (observed in our crystal structures) to a putative activated state as observed in  $FhaC$ , potentially representing the protease-sensitive state observed previously<sup>32</sup>, and the third is the lateral opening event (extracellular surface view).

resulted in considerable upregulation of the quality control protein DegP. Functional analysis of BamA mutants showed that those which supported reduced growth in LB shaking culture still supported LamB expression and trimer formation. Glu717Ala/Asp740Ala-mediated LamB folding, however, appeared reduced (Fig. 3e).

Finally, we studied the biogenesis of these BamA mutants (Fig. 3e, f). When expressed in the presence of endogenous BamA, all mutants except VRGF>A and Asp740Arg showed an indication of folding with heat-modifiable gel-shifts. Proteinase K cleavage products were observed for all mutants except VRGF>A, demonstrating exposure to the extracellular space and thus outer membrane insertion for all other constructs. Taken together, these results suggest that the interaction of Arg 661 with the interior barrel wall is important for proper folding of BamA, and the unique eL6 insertion may have a role in the efficiency of *E. coli* BamA.

## BamA distorts the outer membrane

Compared to other OMPs, the hydrophobic belt of the BamA  $\beta$ -barrel is greatly reduced in width along the C-terminal strand ( $\sim 9$  Å) compared to the opposite side of the barrel ( $\sim 20$  Å) (Supplementary Fig. 13). While proposing that this could destabilize the local membrane environment, we used molecular dynamics simulations to investigate the effects of BamA on membrane stability. Lipid order within the membrane was assessed by looking at the order parameter ( $S_{CD}$ ) for the lipid tails in simulations of NgBamA as well as FhaC and BtuB<sup>34</sup> controls. For NgBamA, lipids close to the C-terminal strand (approximately 15–25 lipids within 12 Å of residue 788) had an approximately threefold decrease in order compared to those along the opposite side of the  $\beta$ -barrel (within 12 Å of residue 531), whereas only a marginal difference was observed for analogous positions on FhaC and BtuB (Fig. 4a and Supplementary Fig. 14). Furthermore, by looking at the mass density of lipid glycerol groups as a measure for membrane thickness, we found that for NgBamA the membrane thickness near the C-terminal strand (centred at residue 788) was 16 Å less than along the opposite side of the barrel, whereas for the reference FhaC, no decrease was seen (Fig. 4b and Supplementary Fig. 15). This marked decrease in lipid order and membrane thickness near the C-terminal strand of NgBamA leads us to suggest that one function of the  $\beta$ -barrel is to prime the membrane for OMP insertion.

Comparison of the HdBamA $\Delta 3$  and NgBamA structures revealed a possible gating mechanism involving POTRA5 that could regulate substrate access to the inside of the  $\beta$ -barrel. To investigate pore access and barrel stability, we monitored the effect of the C-terminal strand on the stability of the  $\beta$ -barrel (Fig. 4c, d and Supplementary Fig. 16). Simulations unexpectedly demonstrated a lateral opening event in the  $\beta$ -barrel of both structures via separation of the first and last  $\beta$ -strands. For HdBamA $\Delta 3$ , whose structure contained an ordered C-terminal strand and POTRA5 oriented away from the  $\beta$ -barrel, the separation between  $\beta$ -strands 1 and 16 ranged from 4 Å (X-ray structure) to 7.5 Å (molecular dynamics simulation), and became more destabilized as the simulation progressed, with the largest openings occurring towards the end of the experiment (Fig. 4e and Supplementary Fig. 16). By contrast, strand separation in NgBamA, which had a disordered C-terminal strand and POTRA5 interacting with the periplasmic loops of the  $\beta$ -barrel, was almost immediate and ranged from 5 Å (X-ray structure) to 10 Å (molecular dynamics simulation), with a larger separation than observed for HdBamA $\Delta 3$  at the same temperature (Fig. 4c–e, Supplementary Fig. 16 and Supplementary Video 1). NgBamA simulation performed at a lower temperature (310 K) exhibited the same lateral opening delayed by approximately 1  $\mu$ s. Control simulations of FhaC and BtuB revealed no such lateral opening at any temperature over the same time scales. Similarly, the average distance between strands 13 and 14 in NgBamA was stable during the simulation (Fig. 4e). Lateral openings in a  $\beta$ -barrel have only been observed in three other structures: FadL<sup>35</sup>, PagP<sup>36</sup> and OmpW<sup>37</sup>, all of which transport hydrophobic molecules. Surprisingly, not only did we observe an opening between  $\beta$ -strands 1 and 16 in the simulations for both HdBamA $\Delta 3$  and NgBamA, but also a closing event such that once opened, the  $\beta$ -barrel did not fully unfold as might be expected. Stabilization of the  $\beta$ -barrel may be attributed to both the intimate interactions between the extracellular loops and specific interactions between eL6 with the opposite side of the barrel via the conserved VRGF/Y motif.

## Concluding remarks

In summary, we find that BamA can perturb the outer membrane in at least two ways: by a reduced hydrophobic surface near  $\beta$ -strand 16 that results in decreased lipid order and membrane thickness, and by transient separation of  $\beta$ -strands 1 and 16 that produces a lateral opening in the barrel. Taken together with movements of the POTRA domains, a highly dynamic membrane environment is created by BamA in the immediate vicinity of the BAM complex. Elegant biophysical analysis of OMP biogenesis *in vitro* has suggested that some  $\beta$ -barrels can be

folded in the periplasm before insertion into the outer membrane<sup>38,39</sup>, but it was unclear how these barrels could insert into the lipid bilayer and whether this is the only folding mechanism used by OMPs. Our studies reveal structural features of BamA that would catalyse the entry of  $\beta$ -barrels into the outer membrane, and we propose two possible mechanisms (Supplementary Video 2). The first mechanism uses the hypothetical conformational switch of loop 6, the POTRA gating motion, and the lateral opening event working together to thread nascent OMPs through the  $\beta$ -barrel of BamA directly into the outer membrane (Fig. 4f). This mechanism would use the exposed strands of BamA (from the lateral opening event) as a template to initiate barrel formation by  $\beta$ -augmentation, forming a transient BamA–OMP complex. The OMP would continue to fold until it eventually buds off from BamA and is released into the outer membrane. More complex OMPs would favour this BamA-assisted approach. However, simpler OMPs that can readily fold into membranes might use a second mechanism independent of the  $\beta$ -barrel domain of BamA. Here, nascent OMPs may be trafficked into close proximity of the outer membrane by interactions with the POTRA domains of BamA for direct insertion into the locally destabilized membrane. Whether similar mechanisms are observed in eukaryotic systems remains to be determined, but our improved model for the mitochondrial homologue Sam50 (also known as Tob55) will assist future studies (Supplementary Figs 17 and 18 and Supplementary Information). Our work represents an important step forward in understanding bacterial OMP membrane insertion, and future investigations will determine whether insight gained from BamA represents a universal mechanism for the biogenesis of all  $\beta$ -barrel membrane proteins.

## METHODS SUMMARY

For overexpression, constructs were prepared using a modified pET20b vector containing an N-terminal 10 $\times$  His tag and TEV protease site. HdBamA $\Delta 3$  and NgBamA were expressed in BL21 (DE3) cells at 20 °C and purified via the N-terminal 10 $\times$  His tag, which was later removed by TEV digestion. Final sample preparation was performed using a size exclusion column and crystallization performed using DMPC:CHAPSO bicelles<sup>40,41</sup>. X-ray data for selenium–single-wavelength anomalous dispersion (Se-SAD) phasing were collected at the I02 beamline at Diamond Light Source, and native data were collected at the GM/CA-CAT and SER-CAT beamlines of the Advanced Photon Source synchrotron at the Argonne National Laboratory. For functional assays, wild-type *E. coli* BamA and mutants were prepared using a modified pRSF-1b vector (EMD Millipore). Mutagenesis and functional assays were performed using JCM-166 cells<sup>13</sup>. Homology modelling for *E. coli* BamA and *Saccharomyces cerevisiae* Sam50 was performed manually using CHAINSAW/CCP4 (refs 42, 43), COOT<sup>44,45</sup> and Chiron<sup>46</sup>. Molecular dynamics simulations were performed using the Anton supercomputer<sup>47</sup>. For more details see Methods.

**Full Methods** and any associated references are available in the online version of the paper.

**Received 21 December 2012; accepted 1 August 2013.**

**Published online 1 September 2013.**

1. Dalbey, R. E., Wang, P. & Kuhn, A. Assembly of bacterial inner membrane proteins. *Annu. Rev. Biochem.* **80**, 161–187 (2011).
2. White, S. H. & von Heijne, G. How translocons select transmembrane helices. *Annu. Rev. Biophys.* **37**, 23–42 (2008).
3. du Plessis, D. J., Nouwen, N. & Driessen, A. J. The Sec translocase. *Biochim. Biophys. Acta* **1808**, 851–865 (2011).
4. Osborne, A. R., Rapoport, T. A. & van den Berg, B. Protein translocation by the SecY1/SecY channel. *Annu. Rev. Cell Dev. Biol.* **21**, 529–550 (2005).
5. Chacinska, A., Koehler, C. M., Milenkovic, D., Lithgow, T. & Pfanner, N. Importing mitochondrial proteins: machineries and mechanisms. *Cell* **138**, 628–644 (2009).
6. Walther, D. M., Rapoport, D. & Tommassen, J. Biogenesis of  $\beta$ -barrel membrane proteins in bacteria and eukaryotes: evolutionary conservation and divergence. *Cell. Mol. Life Sci.* **66**, 2789–2804 (2009).
7. Webb, C. T., Heinz, E. & Lithgow, T. Evolution of the  $\beta$ -barrel assembly machinery. *Trends Microbiol.* **20**, 612–620 (2012).
8. Paschen, S. A., Neupert, W. & Rapoport, D. Biogenesis of  $\beta$ -barrel membrane proteins of mitochondria. *Trends Biochem. Sci.* **30**, 575–582 (2005).
9. Jiang, J. H., Tong, J., Tan, K. S. & Gabriel, K. From evolution to pathogenesis: the link between  $\beta$ -barrel assembly machineries in the outer membrane of mitochondria and Gram-negative bacteria. *Int. J. Mol. Sci.* **13**, 8038–8050 (2012).



10. Tommassen, J. Assembly of outer-membrane proteins in bacteria and mitochondria. *Microbiology* **156**, 2587–2596 (2010).
11. Pugsley, A. P. The complete general secretory pathway in gram-negative bacteria. *Microbiol. Rev.* **57**, 50–108 (1993).
12. Knowles, T. J., Scott-Tucker, A., Overduin, M. & Henderson, I. R. Membrane protein architects: the role of the BAM complex in outer membrane protein assembly. *Nature Rev. Microbiol.* **7**, 206–214 (2009).
13. Wu, T. et al. Identification of a multicomponent complex required for outer membrane biogenesis in *Escherichia coli*. *Cell* **121**, 235–245 (2005).
14. Hagan, C. L., Silhavy, T. J. & Kahne, D.  $\beta$ -Barrel membrane protein assembly by the Bam complex. *Annu. Rev. Biochem.* **80**, 189–210 (2011).
15. Ricci, D. P. & Silhavy, T. J. The Bam machine: a molecular cooper. *Biochim. Biophys. Acta* **1818**, 1067–1084 (2012).
16. Rigel, N. W. & Silhavy, T. J. Making a beta-barrel: assembly of outer membrane proteins in Gram-negative bacteria. *Curr. Opin. Microbiol.* **15**, 189–193 (2012).
17. Hagan, C. L., Kim, S. & Kahne, D. Reconstitution of outer membrane protein assembly from purified components. *Science* **328**, 890–892 (2010).
18. Noinaj, N., Fairman, J. W. & Buchanan, S. K. The crystal structure of BamB suggests interactions with BamA and its role within the BAM complex. *J. Mol. Biol.* **407**, 248–260 (2011).
19. Heuck, A., Schleiffer, A. & Clausen, T. Augmenting  $\beta$ -augmentation: structural basis of how BamB binds BamA and may support folding of outer membrane proteins. *J. Mol. Biol.* **406**, 659–666 (2011).
20. Kim, K. H., Aulakh, S. & Paetzel, M. Crystal structure of  $\beta$ -barrel assembly machinery BamCD protein complex. *J. Biol. Chem.* **286**, 39116–39121 (2011).
21. Albrecht, R. & Zeth, K. Structural basis of outer membrane protein biogenesis in bacteria. *J. Biol. Chem.* **286**, 27792–27803 (2011).
22. Sandoval, C. M., Baker, S. L., Jansen, K., Metzner, S. I. & Sousa, M. C. Crystal structure of BamD: an essential component of the  $\beta$ -barrel assembly machinery of gram-negative bacteria. *J. Mol. Biol.* **409**, 348–357 (2011).
23. Dong, C., Hou, H. F., Yang, X., Shen, Y. Q. & Dong, Y. H. Structure of *Escherichia coli* BamD and its functional implications in outer membrane protein assembly. *Acta Crystallogr. D* **68**, 95–101 (2012).
24. Knowles, T. J. et al. Structure and function of BamE within the outer membrane and the beta-barrel assembly machine. *EMBO Rep.* **12**, 123–128 (2011).
25. Gatzeva-Topalova, P. Z., Walton, T. A. & Sousa, M. C. Crystal structure of YaeT: conformational flexibility and substrate recognition. *Structure* **16**, 1873–1881 (2008).
26. Kim, S. et al. Structure and function of an essential component of the outer membrane protein assembly machine. *Science* **317**, 961–964 (2007).
27. Zhang, H. et al. High-resolution structure of a new crystal form of BamA POTRA4–5 from *Escherichia coli*. *Acta Crystallogr. F* **67**, 734–738 (2011).
28. Gatzeva-Topalova, P. Z., Warner, L. R., Pardi, A. & Sousa, M. C. Structure and flexibility of the complete periplasmic domain of BamA: the protein insertion machine of the outer membrane. *Structure* **18**, 1492–1501 (2010).
29. Knowles, T. J. et al. Fold and function of polypeptide transport-associated domains responsible for delivering unfolded proteins to membranes. *Mol. Microbiol.* **68**, 1216–1227 (2008).
30. Clantin, B. et al. Structure of the membrane protein FhaC: a member of the Omp85–TpsB transporter superfamily. *Science* **317**, 957–961 (2007).
31. Fan, E., Fiedler, S., Jacob-Dubuisson, F. & Muller, M. Two-partner secretion of gram-negative bacteria: a single  $\beta$ -barrel protein enables transport across the outer membrane. *J. Biol. Chem.* **287**, 2591–2599 (2012).
32. Rigel, N. W., Ricci, D. P. & Silhavy, T. J. Conformation-specific labeling of BamA and suppressor analysis suggest a cyclic mechanism for  $\beta$ -barrel assembly in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **110**, 5151–5156 (2013).
33. Leonard-Rivera, M. & Misra, R. Conserved residues of the putative L6 loop of *Escherichia coli* BamA play a critical role in the assembly of  $\beta$ -barrel outer membrane proteins, including that of BamA itself. *J. Bacteriol.* **194**, 4662–4668 (2012).
34. Gumbart, J., Wiener, M. C. & Tajkhorshid, E. Coupling of calcium and substrate binding through loop alignment in the outer-membrane transporter BtuB. *J. Mol. Biol.* **393**, 1129–1142 (2009).
35. Hearn, E. M., Patel, D. R., Lepore, B. W., Indic, M. & van den Berg, B. Transmembrane passage of hydrophobic compounds through a protein channel wall. *Nature* **458**, 367–370 (2009).
36. Cuesta-Seijo, J. A. et al. PagP crystallized from SDS/cosolvent reveals the route for phospholipid access to the hydrocarbon ruler. *Structure* **18**, 1210–1219 (2010).
37. Hong, H., Patel, D. R., Tamm, L. K. & van den Berg, B. The outer membrane protein OmpW forms an eight-stranded beta-barrel with a hydrophobic channel. *J. Biol. Chem.* **281**, 7568–7577 (2006).
38. Tamm, L. K., Hong, H. & Liang, B. Folding and assembly of beta-barrel membrane proteins. *Biochim. Biophys. Acta* **1666**, 250–263 (2004).
39. Tamm, L. K., Arora, A. & Kleinschmidt, J. H. Structure and assembly of beta-barrel membrane proteins. *J. Biol. Chem.* **276**, 32399–32402 (2001).
40. Agah, S. & Faham, S. Crystallization of membrane proteins in bicelles. *Methods Mol. Biol.* **914**, 3–16 (2012).
41. Ujwal, R. & Bowie, J. U. Crystallizing membrane proteins using lipidic bicelles. *Methods* **55**, 337–341 (2011).
42. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
43. Stein, N. CHAINSAW: a program for mutating pdb files used as templates in molecular replacement. *J. Appl. Crystallogr.* **41**, 641–643 (2008).
44. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
45. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
46. Ramachandran, S., Kota, P., Ding, F. & Dokholyan, N. V. Automated minimization of steric clashes in protein structures. *Proteins* **79**, 261–270 (2011).
47. Shaw, D. E. et al. Anton, a special-purpose machine for molecular dynamics simulation. *Commun. ACM* **51**, 91–97 (2008).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank H. Bernstein and R. Ieva for providing JCM-166 cells, J. Beckwith and R. Misra for providing antibodies, and A. M. Stanley and T. Barnard for discussions and comments on the manuscript. N.N., A.J.K., N.C.E., H.C. and S.K.B. are supported by the Intramural Research Program of the National Institutes of Health (NIH), National Institute of Diabetes and Digestive and Kidney Diseases. J.C.G. acknowledges support from the NIH under grants K22-AI100927 and R01-GM67887. P.L. is supported by a postdoctoral fellowship through the Diamond Light Source. T.L. is an Australian Research Council (ARC) Federation Fellow and acknowledges support from ARC Discovery Project (DP120101878) and ARC Linkage International Grant (LX0776170). We thank the respective staffs at the Southeast Regional Collaborative Access Team (SER-CAT) and General Medicine and Cancer Institute's Collaborative Access Team (GM/CA-CAT) beamlines at the Advanced Photon Source, Argonne National Laboratory and the Diamond Light Source for their assistance during data collection. Use of the Advanced Photon Source was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences, under contract no. W-31-109-Eng-38 (SER-CAT), and by the US Department of Energy, Basic Energy Sciences, Office of Science, under contract no. DE-AC02-06CH11357 (GM/CA-CAT). Anton computer time was provided by the National Resource for Biomedical Supercomputing and the Pittsburgh Supercomputing Center through Grant RC2GM093307 from the NIH, using a machine donated by DE Shaw Research.

**Author Contributions** N.N., H.C., N.C.E. and S.K.B. cloned, expressed and purified HdBamA $\Delta$ 3 and NgBamA. P.L. performed data collection for experimental phasing. N.N. crystallized and solved the HdBamA $\Delta$ 3 and NgBamA crystal structures. N.N. and A.J.K. performed the homology modelling and functional assays. J.C.G. designed, conducted and analysed the molecular dynamics simulations. N.N., A.J.K. and S.K.B. analysed and discussed all data. T.L. and S.K.B. conceived and designed the original project. N.N., A.J.K., T.L. and S.K.B. wrote the manuscript.

**Author Information** Atomic coordinates and structure factors for HdBamA $\Delta$ 3 and NgBamA are deposited in the Protein Data Bank under accession codes 4K3C and 4K3B, respectively. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.K.B. ([skbucan@helix.nih.gov](mailto:skbucan@helix.nih.gov)).

## METHODS

**Cloning, expression and purification of *HdBamAΔ3*.** *HdBamAΔ3* from *H. ducreyi* (strain 35000HP/ATCC 700724) was subcloned into a modified pET20b vector (EMD Millipore) containing an N-terminal pelB signal sequence, a 10× His tag and a TEV site starting with residue Tyr262. Expression was performed in BL21(DE3) cells at 20 °C without induction in TB media supplemented with 100 μg ml<sup>-1</sup> carbenicillin. For purification, cells were resuspended in lysis buffer (50 mM Tris-HCl, pH 7.5, 200 mM NaCl, 1 mM MgCl<sub>2</sub>, 10 μg ml<sup>-1</sup> DNaseI, 100 μg ml<sup>-1</sup> 4-(2-aminoethyl)benzenesulphonyl fluoride (AEBSF)) and lysed by two passages through an Emulsiflex C3 (Avestin) homogenizer at 4 °C. The lysate was centrifuged at 12,000g for 10 min to remove unlysed cells, and the supernatant was incubated with 2% Triton X-100 for 30 min at room temperature. The mixture was centrifuged at 160,000g for 90 min at 4 °C. The membrane pellets were resuspended in 50 mM Tris-HCl, pH 7.5, 200 mM NaCl, 20 mM imidazole and solubilized by constant stirring in 5% eluent for 16 h at 4 °C. Solubilized membranes were centrifuged at 265,000g for 60 min at 4 °C, and the supernatant filtered and applied to a 15-ml Ni-NTA column (Qiagen). *HdBamAΔ3* was eluted using 250 mM imidazole. Peak fractions were pooled, concentrated, dialysed and treated with TEV-His protease overnight at 4 °C. To remove uncleaved protein and the TEV-His protease, the protein was re-applied to a second Ni-NTA column, concentrated and then applied to an S-300HR sephacryl size exclusion column (GE Healthcare) using 20 mM Tris-HCl, pH 7.5, 200 mM NaCl, 10 mM sodium citrate, 10 mM lithium sulphate, 0.8% C<sub>6</sub>E<sub>4</sub> and 0.02% Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub>. Peak fractions were verified using SDS-PAGE. Selenomethionine-substituted protein was prepared as previously described<sup>48</sup>.

**Cloning, expression and purification of *NgBamA*.** *NgBamA* from *N. gonorrhoeae* (strain ATCC 700825/FA 1090) was subcloned into the pET20b vector (EMD Millipore) containing an N-terminal pelB signal sequence, a 10× His tag and a TEV site starting with residue Phe23. Expression was performed in BL21(DE3) cells at 20 °C without induction in TB media supplemented with 100 μg ml<sup>-1</sup> carbenicillin. Purification was performed as described for *HdBamAΔ3*.

**Crystallization and data collection.** For crystallization, native *HdBamAΔ3* and *NgBamA* were concentrated to 10 mg ml<sup>-1</sup>, and sparse matrix screening was performed using a TTP Labtech Mosquito crystallization robot using hanging drop vapour diffusion and plates incubated at 21 °C. Although detergent screening, lipidic cubic phase screening and bicelle screening were all performed, only bicelle crystallization<sup>40</sup> produced well diffracting crystals leading to structure determination. Here, 40 μl of protein solution was mixed with 10 μl of 35% bicelle mixture (DMPC:CHAPSO at 2.8:1 ratio)<sup>41</sup>, mixed and incubated on ice for at least 30 min before setting trays. The best native crystals for *HdBamAΔ3* were grown from 100 mM Na-citrate, 100 mM HEPES, pH 7.5, 12% 2-methyl-2,4-pentanediol (MPD). Selenomethionine-substituted crystals of *HdBamAΔ3* were crystallized using the same conditions as for native. The best crystals for *NgBamA* were grown from 0.1 M NaCl, 0.1 M K-phosphate, pH 7.0, 32% PEG 300 and 200 mM Na-malonate. Crystals were collected directly from the crystallization drops and native data were collected at the SER-CAT (ID) and the GM/CA-CAT (ID-D) beamlines of the Advanced Photon Source of the Argonne National Laboratory. Data collection for selenium-single-wavelength anomalous dispersion (Se-SAD) phasing of the *HdBamAΔ3* structure was performed at the I02 beamline of the Diamond Light Source. All data were processed using either HKL2000 or XDS. A summary of the data collection statistics can be found in Supplementary Table 1.

**Structure determination.** For *HdBamAΔ3* and *NgBamA* native data sets, molecular replacement attempts using the known structures of the POTRA domains (PDB codes 2QCZ, 3EFC, 2QDF and 3QB6) and a model for the β-barrel domain using FhaC (PDB code 2QDZ) were unsuccessful. Although we were unable to grow selenomethionine-substituted *NgBamA* in large enough quantities for crystallization, we were able to express selenomethionine-substituted *HdBamAΔ3* and grew crystals using the native conditions. We were able to use these crystals to collect a 2.91 Å SAD data set at the selenium peak wavelength. AutoSol (PHENIX)<sup>49</sup> was then used to locate the selenium sites, finding seven out of eight possible sites and producing a density-modified map that we could build an initial model into. Density for POTRA 4 was largely disordered, explaining why the methionine in this domain was not useful for phasing. However, because POTRA 4 mediates crystal packing, rigid body refinement was used to optimally place this domain, which was built without side chains. We then used this initial structure to solve the native *HdBamAΔ3* structure to 2.9 Å resolution, with *R*/*R*<sub>free</sub> values of 0.22/0.27. The *NgBamA* crystal structure was then solved by molecular replacement in Phaser-MR (PHENIX)<sup>49</sup> using search models based on the β-barrel domain of *HdBamAΔ3* and the reported POTRA domains of *EcBamA* (PDB codes 2QCZ, 3EFC, 2QDF and 3QB6). Here, only the β-barrel domain and POTRA1 and POTRA2 could be placed by molecular replacement (no solutions for POTRA3, POTRA4 or POTRA5). However, POTRA3, POTRA4 and POTRA5 were then manually placed based on weak difference density and subsequent rigid body refinement to optimize the

domain positions. *NgBamA* was solved to a final resolution of 3.2 Å with *R*/*R*<sub>free</sub> values of 0.23/0.28. All model building was performed using COOT<sup>44,45</sup> and subsequent refinement done in PHENIX<sup>49</sup>. R.m.s.d. analysis was performed within PyMOL (Schrödinger). We would like to note that the conformational switch between the POTRA domains of *HdBamAΔ3* (open) and *NgBamA* (closed) is supported by previous studies<sup>50</sup>, however, we cannot exclude this observation being due to a crystallization artefact. Therefore, further studies are needed to validate this hypothesis. For all structures, figures were made with PyMOL (Schrödinger) or Chimera<sup>51</sup> and annotated and finalized with Adobe Illustrator.

**Homology modelling.** For the *BamA* homology model from *E. coli* (*EcBamA*), a pairwise sequence alignment of *BamA* from *E. coli* and *H. ducreyi* was performed using ClustalW<sup>52</sup>. The alignment was then input into CHAINSAW (CCP4)<sup>42,43</sup> along with the structure of *HdBamAΔ3* to produce an initial model of *EcBamA*. Insertions were then added manually within COOT<sup>44,45</sup> and a disulphide bond restraint added within loop 6 between residues Cys 690 and Cys 700 (not present in either *HdBamA* or *NgBamA*). The POTRA domains (from CHAINSAW) were then replaced with the known structures (PDB codes 2QDF and 3QB6) and each POTRA domain aligned according to the *NgBamA* full length structure. Missing side chains were added using Deepview/Swiss-PdbViewer<sup>53</sup> and final model minimization performed using Chiron<sup>46</sup>. For the *Sam50* homology model from *Saccharomyces cerevisiae* (*ScSam50*), a pairwise sequence alignment of the β-barrel of *HdBamA* was performed for all *Sam50* homologues shown in Supplementary Fig. 18 using ClustalW<sup>52</sup> by restricting the alignment to the last ~430 residues, with the most convincing alignment being for *S. cerevisiae* based on an even spread of identities and similarities. The alignment was then fed into CHAINSAW (CCP4)<sup>42,43</sup> along with the structure of *HdBamAΔ3* to produce an initial model of the β-barrel of *ScSam50*. Insertions were then added manually within COOT<sup>44,45</sup>. The N-terminal domain of *ScSam50* was then modelled as a single POTRA domain based on secondary structure predictions (Supplementary Figs 18 and 19), which indicated only a single POTRA domain containing the conserved β-α-α-β fold. Here, the model for the *ScSam50* POTRA domain was built manually within COOT<sup>44,45</sup> and secondary structure elements aligned to POTRA 1 of *EcBamA* to retain the conserved overall fold. The models for the β-barrel and the POTRA domain of *ScSam50* were then aligned to the β-barrel and POTRA 5 of the *NgBamA* crystal structure, merged into a single model, missing side chains added using Deepview/Swiss-PdbViewer<sup>53</sup> and final model minimization performed using Chiron<sup>46</sup>.

**Molecular dynamics simulations.** Systems containing the full-length *NgBamA* and *HdBamA*, that is, with POTRA domains 5 and 2, respectively, were first constructed by placing the protein in a 1,2-dimyristoyl-*sn*-glycero-3-phosphoethanolamine (DMPE) lipid bilayer, used as an outer membrane mimic as done previously<sup>34,54</sup>. The resulting protein-membrane complex was then solvated and K<sup>+</sup> and Cl<sup>-</sup> ions added to a concentration of 150 mM; the *NgBamA* system contained 240,000 atoms and *HdBamA* 206,000 atoms. Equilibration of the systems was carried out in stages for 30 ns using the simulation program NAMD<sup>55</sup>. From these equilibrated systems, new ones were constructed in which the POTRA domains were truncated at residue 417 in *NgBamA* and 419 in *HdBamA*. Such truncation was necessary to fit the limitations of the Anton supercomputer<sup>47</sup>. Anton makes simulations on the time scale of microseconds routinely accessible, and was used for all production simulations reported here. The CHARMM27/CMAP force field for proteins<sup>47,56</sup>, TIP3P for water<sup>57</sup> and CHARMM36 for lipids<sup>58</sup> were used. The simulations were run in the NPT ensemble, thereby allowing the membrane to expand in response to *BamA* opening.

Final system sizes were 110,000 (dimensions of 116 × 113 × 84 Å<sup>3</sup>) and 85,000 (102 × 106 × 79 Å<sup>3</sup>) atoms for *NgBamA* and *HdBamA*, respectively. The number of lipids for each system was 406 (*NgBamA*) and 304 (*HdBamA*), whereas the number of water molecules was 20,000 (*NgBamA*) and 15,000 (*HdBamA*). Proteins were at least 50–60 Å from their periodic images in the membrane plane and 15 Å along the normal axis. As controls, FhaC and the vitamin B<sub>12</sub> transporter BtuB were also simulated using the same parameters as the *BamA* systems. The POTRA domains of FhaC were removed as done for *BamA*, such that the N-terminal helix (residues 1–30) and the barrel (residues 207–554) were retained. The missing extracellular loop (residues 384–397) was modelled and inserted into the structure. The size of the final FhaC system was 83,000 atoms, with dimensions 92 × 94 × 96 Å<sup>3</sup>; for BtuB the size was 71,000 atoms (83 × 81 × 103 Å<sup>3</sup>). The FhaC and BtuB systems contained 251 and 173 lipids and 16,000 and 14,000 water molecules, respectively.

Two temperatures were used for *NgBamA* (310 K and 340 K), FhaC (310 K and 340 K) and BtuB (325 K and 353 K) for two reasons: first, each temperature permits a different phase of the membrane (gel-like at 310 K and fluid at 340 K) to be explored, and second, the probability of fluctuations in the barrel opening are expected to be increased at a higher temperature. Temperatures of 353–490 K have previously been validated for peptide-membrane partitioning studies and were found to not considerably affect the systems' thermodynamic properties<sup>59</sup>. Regardless of

the temperature, the FhaC and BtuB barrels were found to be stable, whereas NgBamA was not. HdBamA was apparently more stable than NgBamA at 340 K, although barrel opening was nonetheless observed after approximately 1  $\mu$ s. Two temperatures were used for NgBamA (310 K and 340 K), FhaC (310 K and 340 K) and BtuB (325 K and 353 K) for two reasons. First, each temperature permits a different phase of the membrane (gel-like at 310 K and fluid at 340 K) to be explored, and second, the probability of fluctuations in the putative lateral opening of the  $\beta$ -barrel are expected to be increased at a higher temperature. We note that the gel-like membrane at 310 K is probably more representative of the expected low fluidity for the outer membrane. The total time for all Anton simulations is 7.5  $\mu$ s.

**Functional assessment of EcBamA mutants.** For functional analysis of mutations of EcBamA, the coding sequence for EcBamA containing an N-terminal pelB signal sequence and 10 $\times$  His tag was placed into the pRSF-1b vector (EMD Millipore). Mutants were prepared using standard site-directed mutagenesis protocols (primer sequences available on request). JCM-166 cells, whose endogenous BamA is under the control of an arabinose promoter<sup>13</sup>, were transformed via electroporation with pRSF1 vectors coding for a mutant BamA under kanamycin selection, and were then plated on LB+ agar plates with 50  $\mu$ g ml<sup>-1</sup> kanamycin in the presence or absence of 0.1% L-(+)-arabinose and grown for 12–15 h at 37 °C.

BamA mutants that exhibited decreased colony formation compared to wild type were further analysed in culture growth assays. Colonies on a 0.1% arabinose plate were isolated and grown in LB media plus 50  $\mu$ g ml<sup>-1</sup> kanamycin at 37 °C. The expression of endogenous BamA on the arabinose-containing plates allowed at least some initial growth of all mutant transformants, and after 3–4 h all cultures were washed in PBS and diluted in fresh LB plus kanamycin to an absorbance ( $A_{600\text{ nm}}$ ) of  $\sim$ 0.04. Cultures were incubated shaking at 37 °C, with small aliquots taken for triplicate  $A_{595\text{ nm}}$  measurements on a BioRad iMark Microplate Reader every hour. Preliminary experiments demonstrated that those mutants that grew similar to wild type on plates also grew similar to wild type in growth curves. Additional images from our colony growth assays and preliminary growth curves are available on request.

For western blot analysis of mutant BamA and DegP expression levels, colonies grown on LB agar plus kanamycin and 0.1% arabinose were isolated and grown in the absence of arabinose in M63 minimal media<sup>60</sup> with 0.4% glycerol, 0.1% casamino acids and 1 mM MgSO<sub>4</sub> supplements plus 50  $\mu$ g ml<sup>-1</sup> kanamycin at 30 °C. Cells were collected at an  $A_{600\text{ nm}}$  of 0.5–1.0, boiled at 99 °C for 10 min and lysates were resolved with NuPAGE (Invitrogen). Antibodies used were: anti-His-horseradish peroxidase (HRP) conjugate, 1:6,000 (Sigma); anti-LamB, 1:2,000 (provided by R. Misra); anti-DegP, 1:5,000 (provided by J. Beckwith); anti-GroEL, 1:20,000 (Sigma); anti-MBP, 1:5,000 (NEB); anti-mouse HRP, 1:15,000 (Sigma); and anti-rabbit HRP, 1:15,000 (Sigma). Antibody staining was visualized with the ECL prime kit (GE Healthcare).

**Heat-modifiability assays.** To determine whether the BamA mutants were folded properly, heat-modifiability assays were performed using whole cells or whole-cell lysates. Transformed cells were cultured overnight at 30 °C in LB plus 0.1% arabinose, permitting folding and outer membrane incorporation by the endogenous BamA. To prepare whole-cell lysates, cells were centrifuged, resuspended in 1 $\times$  PBS, 1%  $\beta$ -dodecylmaltoside (DDM), 10 mM EDTA, 10  $\mu$ g ml<sup>-1</sup> lysozyme, supplemented with AEBSF and DNaseI, and rocked at room temperature for 15 min before being centrifuged again for 10 min at 21,000g using a microcentrifuge. SDS-loading

buffer was then added to either whole cells (an estimated  $2 \times 10^7$  cells) or lysates, and either boiled (0.5% or 1% SDS in loading buffer) or left at room temperature (0.1% SDS in loading buffer) for 5–10 min. The samples were then separated using NativePage 4–12% gels (Invitrogen) with 1 $\times$  MES SDS-PAGE running buffer by running the gels for 60 min at 150 V (constant) at 4 °C. Transfer to PVDF (polyvinylidene difluoride) membrane was performed using the iBlot system (Invitrogen) and anti-His-HRP (Sigma) antibodies were used for western blot analysis and imaged using an ImageQuant LAS 4000 imaging system (GE Healthcare).

**Proteinase K digestion assays.** To access the surface exposure of BamA and mutants, proteinase K digestion assays were performed on whole cells as has been reported previously<sup>32</sup>. In brief, colonies on an LB plus 0.1% arabinose agar plate were grown in LB plus 0.1% arabinose liquid culture overnight at 30 °C. Cells (0.5–1.0 ml) at  $A_{600\text{ nm}}$  of 1.0 were centrifuged and resuspended in 20 mM Tris-HCl, pH 7.4, 100 mM NaCl. Then, 0.5 mg ml<sup>-1</sup> proteinase K was added and incubated at 37 °C for 30 min. Phenylmethanesulfonylfluoride (PMSF) was then added to a final concentration of 10 mM and the cells were centrifuged again and supernatant removed. PMSF was then added again and then cells resuspended in 1 $\times$  LDS loading buffer, boiled and separated using NuPAGE analysis. Transfer to PVDF membrane was performed using the iBlot system (Invitrogen) and anti-His-HRP (Sigma) and anti-MBP (NEB) antibodies were used for western blot analysis and imaged using an ImageQuant LAS 4000 imaging system (GE Healthcare).

48. Fairman, J. W. *et al.* Crystal structures of the outer membrane domain of intimin and invasins from enterohemorrhagic *E. coli* and enteropathogenic *Y. pseudotuberculosis*. *Structure* **20**, 1233–1243 (2012).
49. Adams, P. D. *et al.* PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D* **58**, 1948–1954 (2002).
50. Robert, V. *et al.* Assembly factor Omp85 recognizes its outer membrane protein substrates by a species-specific C-terminal motif. *PLoS Biol.* **4**, e377 (2006).
51. Pettersen, E. F. *et al.* UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
52. Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).
53. Guex, N. & Peitsch, M. C. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**, 2714–2723 (1997).
54. Barnard, T. J. *et al.* Molecular basis for the activation of a catalytic asparagine residue in a self-cleaving bacterial autotransporter. *J. Mol. Biol.* **415**, 128–142 (2012).
55. Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).
56. Mackerell, A. D., Jr, Feig, M. & Brooks, C. L. III. Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem.* **25**, 1400–1415 (2004).
57. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
58. Klauda, J. B. *et al.* Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J. Phys. Chem. B* **114**, 7830–7843 (2010).
59. Ulmschneider, J. P., Smith, J. C., White, S. H. & Ulmschneider, M. B. *In silico* partitioning and transmembrane insertion of hydrophobic peptides under equilibrium conditions. *J. Am. Chem. Soc.* **133**, 15487–15495 (2011).
60. Elbing, K. & Brent, R. Media preparation and bacteriological tools. *Curr. Protoc. Mol. Biol.* **Chapter 1**, Unit 1.1 (2002).



# A strong magnetic field around the supermassive black hole at the centre of the Galaxy

R. P. Eatough<sup>1</sup>, H. Falcke<sup>1,2,3</sup>, R. Karuppusamy<sup>1</sup>, K. J. Lee<sup>1</sup>, D. J. Champion<sup>1</sup>, E. F. Keane<sup>4</sup>, G. Desvignes<sup>1</sup>, D. H. F. M. Schnitzeler<sup>1</sup>, L. G. Spitler<sup>1</sup>, M. Kramer<sup>1,4</sup>, B. Klein<sup>1,5</sup>, C. Bassa<sup>4</sup>, G. C. Bower<sup>6</sup>, A. Brunthaler<sup>1</sup>, I. Cognard<sup>7,8</sup>, A. T. Deller<sup>3</sup>, P. B. Demorest<sup>9</sup>, P. C. C. Freire<sup>1</sup>, A. Kraus<sup>1</sup>, A. G. Lyne<sup>4</sup>, A. Noutsos<sup>1</sup>, B. Stappers<sup>4</sup> & N. Wex<sup>1</sup>

Earth's nearest candidate supermassive black hole lies at the centre of the Milky Way<sup>1</sup>. Its electromagnetic emission is thought to be powered by radiatively inefficient accretion of gas from its environment<sup>2</sup>, which is a standard mode of energy supply for most galactic nuclei. X-ray measurements have already resolved a tenuous hot gas component from which the black hole can be fed<sup>3</sup>. The magnetization of the gas, however, which is a crucial parameter determining the structure of the accretion flow, remains unknown. Strong magnetic fields can influence the dynamics of accretion, remove angular momentum from the infalling gas<sup>4</sup>, expel matter through relativistic jets<sup>5</sup> and lead to synchrotron emission such as that previously observed<sup>6–8</sup>. Here we report multi-frequency radio measurements of a newly discovered pulsar close to the Galactic Centre<sup>9–12</sup> and show that the pulsar's unusually large Faraday rotation (the rotation of the plane of polarization of the emission in the presence of an external magnetic field) indicates that there is a dynamically important magnetic field near the black hole. If this field is accreted down to the event horizon it provides enough magnetic flux to explain the observed emission—from radio to X-ray wavelengths—from the black hole.

Linearly polarized radio waves that pass through a magnetized medium experience Faraday rotation. The resulting rotation of the polarization vector is given by  $\Delta\phi = \text{RM}\lambda^2$ , where the rotation measure,  $\text{RM} = e^3 / (2\pi m_e^2 c^4) \int B(s)n(s)ds$ , depends on the line-of-sight magnetic field,  $B$ ; the free-electron density,  $n$ ; the path length,  $s$ ; the electron charge,  $e$ , and mass,  $m_e$ ; and the speed of light,  $c$ . The radio emission associated with the Galactic Centre black hole, Sagittarius A\* (Sgr A\*), has  $\text{RM} = -5 \times 10^5 \text{ rad m}^{-2}$ , which is the highest known RM of any source in the Galaxy, and is believed to be due to a column of hot, magnetized gas from the accretion flow onto the black hole<sup>13,14</sup>.

The radio emission from Sgr A\*, however, probes only the innermost scales of accretion. For most accretion models<sup>14</sup>, the term  $B(r)n(r)$  decays much faster than  $r^{-1}$ , where  $r$  is the radial distance from the black hole. Consequently, the Faraday rotation imprinted onto the radio emission from Sgr A\*, which has to pass through the entire column of accreting gas, is dominated by the smallest scales. To measure the magnetization of the accretion flow on the outermost scales, other polarized radio sources, such as pulsars, are needed. A pulsar closely orbiting Sgr A\* would also be an unparalleled tool for testing the spacetime structure around the black hole<sup>15</sup>. Despite predictions that there are more than a thousand pulsars in the central parsec of the Galaxy<sup>16</sup>, there has been a surprising lack of detections<sup>17</sup>, potentially owing to severe interstellar dispersion and scattering in the inner Galaxy<sup>18</sup>.

Recently, the NASA Swift X-ray Telescope detected a bright X-ray flare<sup>9</sup> near Sgr A\* (projected offset of  $\sim 3'' = 0.12 \text{ pc}$  (ref. 19) at a Galactic Centre distance of  $d = 8.3 \text{ kpc}$ ). Subsequent X-ray observations by the NASA NuSTAR telescope resulted in the detection of pulsations with a period of 3.76 s (ref. 10). This behaviour is indicative

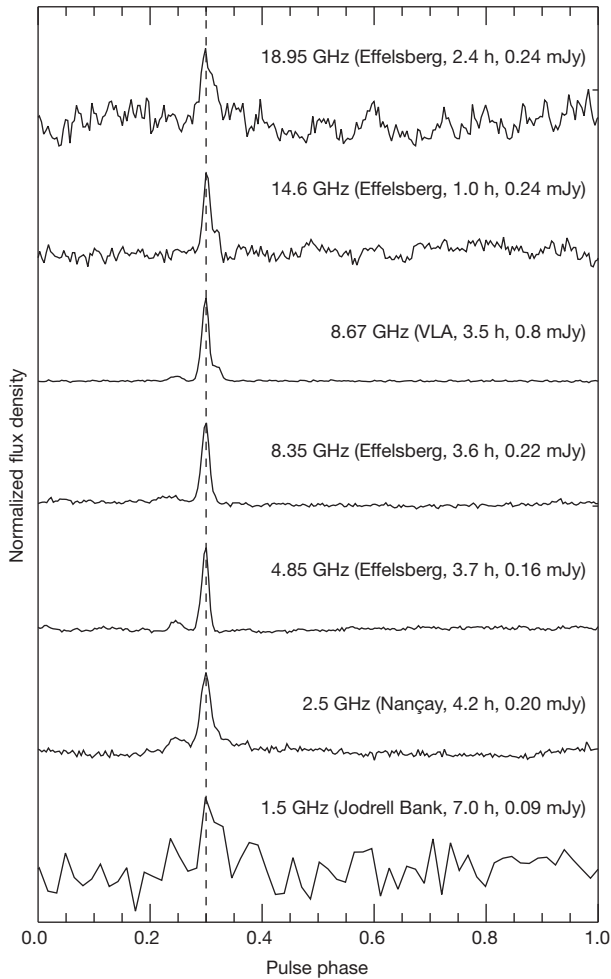
of a magnetar, a highly magnetized pulsar, in outburst. During radio follow-up observations at the MPIfR Effelsberg Radio Observatory on 28 April 2013, the first weak detection of pulsations, with spin parameters matching those reported by NuSTAR, was made. Since then, the pulsar, PSR J1745–2900, has been consistently detected at Effelsberg, at the Paris Observatory-Nançay Radio Astronomy Facility, at the NRAO Karl G. Jansky Very Large Array (VLA), tentatively at The University of Manchester Jodrell Bank Observatory (Fig. 1) and with the CSIRO Australia Telescope Compact Array<sup>12</sup>. Measurements of the delay in the arrival times of pulses at lower frequencies (2.5 GHz) with respect to those at higher frequencies (8.35 GHz) yield an integrated column density of free electrons, the dispersion measure, of  $\text{DM} = 1,778 \pm 3 \text{ cm}^{-3} \text{ pc}$ , which is the highest value measured for any known pulsar. This is consistent with a source located within  $<10 \text{ pc}$  of the Galactic Centre, in the framework of the NE2001 free-electron density model of the Galaxy<sup>20</sup>. Including this source, only four radio-emitting magnetars are known<sup>21</sup> in the Milky Way, making a chance alignment unlikely. If we consider a uniform source distribution occupying a cylinder of radius 10 kpc and height 1 kpc, then the fraction of sources present within an angular distance of  $\sim 3''$  around Sgr A\* is  $\sim 3 \times 10^{-9}$ . Given the current population of radio pulsars ( $\sim 2,000$ ) and radio magnetars, the numbers present within the same region by chance will be  $\sim 6 \times 10^{-6}$  and  $\sim 1 \times 10^{-8}$ , respectively.

The emission from the pulsar is highly linearly polarized<sup>12,22</sup> (Fig. 2). Using the RM synthesis method<sup>23</sup> and measuring the Faraday rotation in three frequency bands and at three different telescope sites, we derive a RM of  $(-6.696 \pm 0.005) \times 10^4 \text{ rad m}^{-2}$  (Fig. 3). This measurement is consistent with that reported elsewhere<sup>12</sup>. The RM is the largest measured for any Galactic object other than Sgr A\*<sup>13,14</sup>, and is more than an order of magnitude larger than all the other RMs measured to within tens of parsecs of Sgr A\*<sup>24</sup>. The RM is also more than what can be optimistically expected as foreground<sup>25</sup>. This constrains the magnetized plasma causing the Faraday rotation (the Faraday screen) to be within some ten parsecs from the Galactic Centre.

A frequently used estimate of the magnetic field is  $B \geq \text{RM} / 0.81 \text{ DM } \mu\text{G}$ , which gives  $B \geq 50 \mu\text{G}$  (ref. 12). However, this is not a stringent limit, because DM and RM are dominated by very different scales. Hence, the extra information about the gas in the central 10 pc must be used for a more robust estimate of the magnetic field.

Two ionized gas phases in the Galactic Centre interstellar medium towards the line of sight of the pulsar could be associated with the Faraday screen: a warm component from the northern arm of the gas streamer Sgr A West<sup>26</sup>, which passes behind Sgr A\*, and a diffuse hot component seen in the X-ray emission<sup>3</sup> with  $T = 2.2 \times 10^7 \text{ K}$ . The warm gas in the northern arm has a width of  $>0.1 \text{ pc}$ , electron densities of  $\sim 10^5 \text{ cm}^{-3}$  measured from radio recombination lines<sup>26</sup>, and a magnetic field of  $\sim 2 \text{ mG}$  (ref. 27). The inferred RM and DM values for

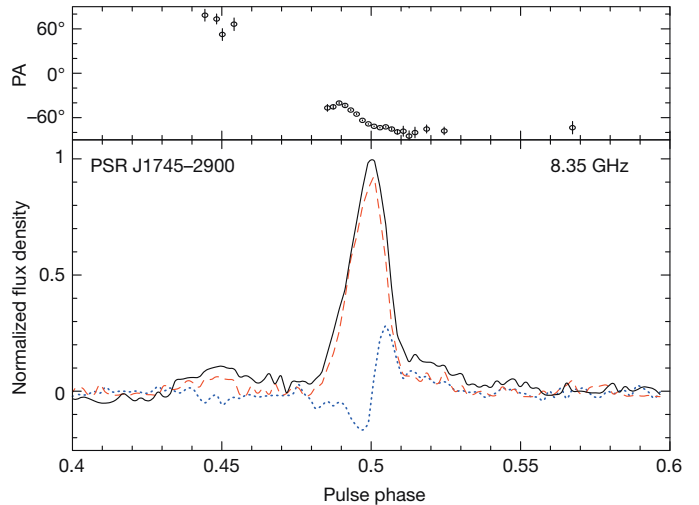
<sup>1</sup>Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, D-53121 Bonn, Germany. <sup>2</sup>Department of Astrophysics, Institute for Mathematics, Astrophysics and Particle Physics, Radboud University, PO Box 9010, 6500 GL Nijmegen, The Netherlands. <sup>3</sup>ASTRON, PO Box 2, 7990 AA Dwingeloo, The Netherlands. <sup>4</sup>Jodrell Bank Centre for Astrophysics, School of Physics and Astronomy, The University of Manchester, Manchester M13 9PL, UK. <sup>5</sup>Bonn-Rhein-Sieg University of Applied Sciences, Grantham-Allee 20, D-53757 Sankt Augustin, Germany. <sup>6</sup>Astronomy Department, B-20 Hearst Field Annex, University of California, Berkeley, California 94720-3411, USA. <sup>7</sup>LPC2E/CNRS - Université d'Orléans, 45071 Orléans, France. <sup>8</sup>Nançay/Paris Observatory, 18330 Nançay, France. <sup>9</sup>National Radio Astronomy Observatory, 520 Edgemont Road, Charlottesville, Virginia 22903, USA.



**Figure 1 | Average pulse profiles of PSR J1745–2900 at each of the radio frequencies where detections have been made.** All observations have been centred on the X-ray position measured with NASA’s Chandra X-ray Observatory<sup>19</sup>. The telescope used, the total observation time required to generate the profile and the average flux density are indicated in brackets after the frequency label. In each case, the profile has been down-sampled from the original sampling interval to 256 phase bins (64 for the Jodrell Bank data), and the peak flux density has been normalized to unity. The profiles have been aligned on the peak of the main pulse detected. By measuring accurate pulse arrival times, we have constructed a coherent timing solution, that is, a model that tracks every single rotation of the pulsar. Between modified Julian dates 56414 and 56426, this model has given a value for the spin period of  $P = 3.76354676(2)$  s and a value for the time derivative of the period (spin-down) of  $\dot{P} = 6.82(3) \times 10^{-12}$ ; uncertainties in the last digit, given in brackets, are derived from the  $1\sigma$  error of the timing model fit. Absolute timing from 1.5 to 8.35 GHz has established that the main pulse in each profile is indeed aligned at each frequency.

a source in or behind the northern arm are  $RM \approx 2 \times 10^7 \text{ rad m}^{-2}$  (for an ordered magnetic field) and  $DM \approx 10^4 \text{ pc cm}^{-3}$ . The measured DM and RM values therefore place the pulsar and the screen in front of the northern arm<sup>26</sup>.

Consequently, the Faraday screen must be associated with the hot gas component, for which no magnetic field estimates yet exist. The density in the hot gas shows a radial fall-off as a function of  $r$ . At 0.4 pc ( $10''$ ) we find that  $n \approx 26 \text{ cm}^{-3}$ , whereas at 0.06 pc ( $1.5''$ ) it can be inferred that  $n \lesssim 160 \text{ cm}^{-3}$ , using the optically thin thermal plasma model<sup>3</sup>. Farther away, on the 40-pc scale<sup>28</sup> ( $17''$ ), the density has decreased to 0.1–0.5  $\text{cm}^{-3}$  and we can roughly describe the density within the central parsecs with a profile of the form  $n(r) \approx 26 \text{ cm}^{-3} (r/0.4 \text{ pc})^{-1}$ . The contribution of this hot gas component to DM is of order  $10^2 \text{ cm}^{-3} \text{ pc}$ . This is consistent with the modest increase in DM with respect to the hitherto closest known pulsars to the Galactic Centre.



**Figure 2 | Pulse profile of PSR J1745–2900 at 8.35 GHz.** After correcting for the Faraday rotation of  $(-6.696 \pm 0.005) \times 10^4 \text{ rad m}^{-2}$ , we can measure the intrinsic polarization across the pulse profile, together with the polarization position angle (PA). The degree of linear polarization (red dashed line) is nearly 100%, and a significant amount ( $\sim 15\%$ ) of circular polarization (blue dotted line) is also detected. A consistent ‘S’-shaped PA swing is measured at each frequency.

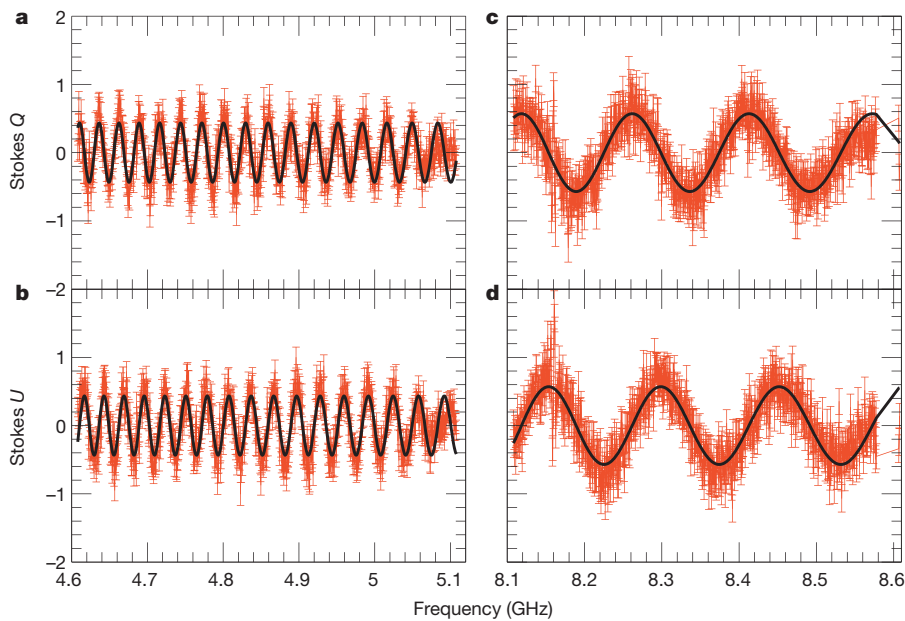
For a simple one-zone Faraday screen, where  $RM \propto B(r)n(r)r$ , we have  $RM = 8.1 \times 10^5 B(r)n(r)r \text{ rad m}^{-2}$ , where  $B(r)$  is in units of Gauss,  $n(r)$  is expressed in units of  $\text{cm}^{-3}$  and  $r$  is expressed in parsecs. Using the density prescription above with an  $r^{-1}$  scaling, we find that  $B \gtrsim 8[RM/(66,960 \text{ m}^{-2})][n_0/(26 \text{ cm}^{-3})]^{-1} \text{ mG}$ . This is a lower limit, because possible turbulent field components or field reversals reduce RM. We note again that this RM value is indeed dominated by the smallest distance scale, that is, by the gas on scales of the de-projected distance,  $r > 0.12 \text{ pc}$ , of the pulsar from Sgr A\*.

This  $B$  value is higher than the magnetic field in the northern arm and is also higher than the equipartition field in the hot phase at this scale. To bring thermal and magnetic energy into equipartition, the gas density at  $r \approx 0.12 \text{ pc}$  would need to increase by a factor of three, to  $260 \text{ cm}^{-3}$ , yielding  $B \approx 2.6 \text{ mG}$ . If there were many field reversals within the Faraday screen, the magnetic field would be driven to values much greater than the equipartition field, suggesting that a relatively ordered magnetic field is pervading the hot gas close to the super-massive black hole.

Because Sgr A\* accretes from this magnetized hot phase, density and magnetic field will further increase at smaller radii. Emission models of Sgr A\* require magnetic fields of about 30–100 G to explain the synchrotron radiation from near the event horizon<sup>6–8</sup>. Hence, if the gas falls from  $3 \times 10^5$  Schwarzschild radii (0.12 pc) down to a few Schwarzschild radii, a simple  $B \propto r^{-1}$  scaling would be enough to provide a magnetic field of several hundred gauss. This is well within the range of most accretion models, where equipartition between magnetic, kinetic and gravitational energy in the accreting gas is assumed<sup>14,29</sup>.

The field at large radius in the accretion flow onto Sgr A\* is therefore sufficient to provide the necessary field at small radius, via simple accretion. Moreover, the availability of ordered magnetic fields would make the proposed formation of a jet-like outflow in Sgr A\*<sup>30</sup> viable. Super-equipartition magnetic fields could also suppress accretion and help to explain the low accretion rate of Sgr A\*.

At its projected distance, PSR J1745–2900 could move (owing to orbital motion) through the hot gas surrounding Sgr A\* at several milliarcseconds per year and reveal RM variations as well as proper motion. Continued pulsar polarimetry and very-long-baseline interferometry astrometry can readily measure these effects. Also, given that magnetars constitute only a small fraction of the pulsar population and the excess DM towards the Galactic Centre is not too large, we expect



**Figure 3 | RM synthesis analysis for the radio polarization of PSR J1745–2900.** The red points, with  $1\sigma$  error bars given by the off-pulse, baseline root mean squared value of the polarization profile, show the observed polarized flux density in the Stokes parameters  $Q$  and  $U$ . We note that polarization measurements were not possible at all frequencies, owing to hardware limitations. RM is measured by a two-step method. First we perform the Fourier transformation of the polarization intensity to get the RM Faraday spectrum, the peak of which is used to find a rough estimate of RM. Using this initial value, we then perform a least-squares fit to the  $Q$  and  $U$  curves to find RM and its error. The black curves are the model values based on the best-fit RM. The sinusoidal variation in  $Q$  and  $U$  due to Faraday rotation is seen across the frequency bands centred at 4.85 GHz (a and b) and 8.35 GHz (c and d). At 2.5 GHz (not shown), the variation is so severe that this signature is more easily seen in the RM spectrum. The RM values derived for each frequency band are independently consistent: at 2.5 GHz,  $RM = (-6.70 \pm 0.01) \times 10^4 \text{ rad m}^{-2}$ ; at 4.85 GHz,  $RM = (-6.694 \pm 0.006) \times 10^4 \text{ rad m}^{-2}$ ; and at 8.35 GHz,  $RM = (-6.68 \pm 0.04) \times 10^4 \text{ rad m}^{-2}$ . RM has also been measured with the VLA at 8.67 GHz, giving  $(-6.70 \pm 0.04) \times 10^4 \text{ rad m}^{-2}$ . The combined and appropriately weighted average is  $(-6.696 \pm 0.005) \times 10^4 \text{ rad m}^{-2}$ .

there to be additional observable radio pulsars in the same region. Such pulsars could be used to map out the accretion region around the black hole in more detail, and even to test its space-time properties.

**Online Content** Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 30 May; accepted 24 July 2013.

Published online 14 August 2013.

- Genzel, R., Eisenhauer, F. & Gillessen, S. The Galactic Center massive black hole and nuclear star cluster. *Rev. Mod. Phys.* **82**, 3121–3195 (2010).
- Narayan, R. & Yi, I. Advection-dominated accretion: a self-similar solution. *Astrophys. J.* **428**, L13–L16 (1994).
- Baganoff, F. K. *et al.* Chandra X-ray spectroscopic imaging of Sagittarius A\* and the central parsec of the Galaxy. *Astrophys. J.* **591**, 891–915 (2003).
- Balbus, S. A. & Hawley, J. F. A powerful local shear instability in weakly magnetized disks. I - Linear analysis. II - Nonlinear evolution. *Astrophys. J.* **376**, 214–233 (1991).
- Beckwith, K., Hawley, J. F. & Krolik, J. H. The influence of magnetic field geometry on the evolution of black hole accretion flows: similar disks, drastically different jets. *Astrophys. J.* **678**, 1180–1199 (2008).
- Falcke, H. & Markoff, S. The jet model for Sgr A\*: radio and X-ray spectrum. *Astron. Astrophys.* **362**, 113–118 (2000).
- Mościbrodzka, M., Gammie, C. F., Dolence, J. C., Shiokawa, H. & Leung, P. K. Radiative Models of Sgr A\* from GRMHD simulations. *Astrophys. J.* **706**, 497–507 (2009).
- Dexter, J., Agol, E., Fragile, P. C. & McKinney, J. C. The submillimeter bump in Sgr A\* from relativistic MHD simulations. *Astrophys. J.* **717**, 1092–1104 (2010).
- Kennea, J. A. *et al.* Swift Discovery of a new soft gamma repeater, SGR J1745–29, near Sagittarius A\*. *Astrophys. J.* **770**, L24 (2013).
- Mori, K. *et al.* NuSTAR discovery of a 3.76 s transient magnetar near Sagittarius A\*. *Astrophys. J.* **770**, L23 (2013).
- Eatough, R. P. *et al.* Detection of radio pulsations from the direction of the NuSTAR 3.76 second X-ray pulsar at 8.35 GHz. *Astron. Telegr.* **5040**, 1 (2013).
- Shannon, R. M. & Johnston, S. Radio properties of the magnetar near Sagittarius A\* from observations with the Australia Telescope Compact Array. Preprint at <http://arxiv.org/abs/1305.3036> (2013).
- Bower, G. C., Falcke, H., Wright, M. C. & Backer, D. C. Variable linear polarization from Sagittarius A\*: evidence of a hot turbulent accretion flow. *Astrophys. J.* **618**, L29–L32 (2005).
- Marrone, D. P., Moran, J. M., Zhao, J.-H. & Rao, R. An unambiguous detection of Faraday rotation in Sagittarius A\*. *Astrophys. J.* **654**, L57–L60 (2007).
- Liu, K., Wex, N., Kramer, M., Cordes, J. M. & Lazio, T. J. W. Prospects for probing the spacetime of Sgr A\* with pulsars. *Astrophys. J.* **747**, 1 (2012).
- Wharton, R. S., Chatterjee, S., Cordes, J. M., Deneva, J. S. & Lazio, T. J. W. Multiwavelength constraints on pulsar populations in the Galactic Center. *Astrophys. J.* **753**, 108 (2012).
- Eatough, R. P. *et al.* in *Neutron Stars and Pulsars: Challenges and Opportunities After 80 Years* (ed. Leeuwen, J. V.) 382–384 (Cambridge Univ. Press, 2013).
- Lazio, T. J. W. & Cordes, J. M. Hyperstrong radio-wave scattering in the Galactic Center. II. A likelihood analysis of free electrons in the Galactic Center. *Astrophys. J.* **505**, 715–731 (1998).
- Rea, N. *et al.* Chandra localization of the soft gamma repeater in the Galactic Center region. *Astron. Telegr.* **5032**, 1 (2013).
- Cordes, J. M. & Lazio, T. J. W. NE2001.1. A new model for the galactic distribution of free electrons and its fluctuations. Preprint at <http://arxiv.org/abs/astro-ph/0207156> (2002).
- Lazarus, P., Kaspi, V. M., Champion, D. J., Hessels, J. W. T. & Dib, R. Constraining radio emission from magnetars. *Astrophys. J.* **744**, 97 (2012).
- Lee, K. J. *et al.* Polarisation profiles and rotation measure of PSR J1745–2900 measured at Effelsberg. *Astron. Telegr.* **5064**, 1 (2013).
- Brentjens, M. A. & de Bruyn, A. G. Faraday rotation measure synthesis. *Astron. Astrophys.* **441**, 1217–1228 (2005).
- Law, C. J., Brentjens, M. A. & Novak, G. A constraint on the organization of the Galactic Center magnetic field using Faraday rotation. *Astrophys. J.* **731**, 36 (2011).
- Bower, G. C., Backer, D. C., Zhao, J.-H., Goss, M. & Falcke, H. The linear polarization of Sagittarius A\*. I. VLA spectropolarimetry at 4.8 and 8.4 GHz. *Astrophys. J.* **521**, 582–586 (1999).
- Zhao, J.-H. *et al.* The high-density ionized gas in the central parsec of the Galaxy. *Astrophys. J.* **723**, 1097–1109 (2010).
- Plante, R. L., Lo, K. Y. & Crutcher, R. M. The magnetic fields in the galactic center: detection of H I Zeeman splitting. *Astrophys. J.* **445**, L113–L116 (1995).
- Muno, M. P. *et al.* Diffuse X-ray emission in a deep Chandra image of the Galactic Center. *Astrophys. J.* **613**, 326–342 (2004).
- Macquart, J.-P., Bower, G. C., Wright, M. C. H., Backer, D. C. & Falcke, H. The rotation measure and 3.5 millimeter polarization of Sagittarius A\*. *Astrophys. J.* **646**, L111–L114 (2006).
- Falcke, H., Mannheim, K. & Biermann, P. L. The galactic center radio jet. *Astron. Astrophys.* **278**, L1–L4 (1993).

**Acknowledgements** We wish to thank D. D. Xu, P. Lazarus and L. Guillemot for discussions. We also thank O. Wucknitz and R. Beck for reading the manuscript. R.K., L.G.S. and P.C.C.F. gratefully acknowledge financial support from the European Research Council for the ERC Starting Grant BEACON under contract no. 279702. K.J.L. was funded by ERC Advanced Grant LEAP under contract no. 227947. H.F. acknowledges funding from an Advanced Grant of the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC Grant Agreement no. 227610. This work was based on observations with the 100-m telescope of the MPIfR (Max-Planck-Institut für Radioastronomie) at Effelsberg. The Nançay radio telescope is part of the Paris Observatory, associated with the Centre National de la Recherche Scientifique (CNRS), and partly supported by the Région Centre in France. The National Radio Astronomy Observatory (NRAO) is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc.

**Author Contributions** R.P.E.: initial detections, observations performed at Effelsberg and data processing; H.F.: observational and theoretical background and paper



formulation; R.K.: observational technical assistance and pulsar timing; K.J.L.: polarization and RM measurements; D.J.C.: pulsar timing solution; E.F.K.: flux density calculations, observational assistance and observations at Jodrell Bank; G.D.: observations at Nançay; D.H.F.M.S.: observational background and RM interpretation; L.G.S.: observational background and data processing and analysis; M.K.: observational background and RM interpretation; B.K.: technical observational assistance at Effelsberg; C.B.: observations at Jodrell Bank; G.C.B.: observations at the VLA and RM interpretation; A.B.: observations at the VLA; I.C.: observations at Nançay; A.T.D.: observations at the VLA; P.B.D.: observations at the VLA; P.C.C.F.: observational

background and pulsar timing; A.K.: technical observational assistance at Effelsberg; A.G.L.: observations at Jodrell Bank and help with initial detections; A.N.: observational background and RM interpretation; B.S.: observations at Jodrell Bank; N.W.: theoretical background and orbital characteristics.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to R.P.E. ([reatough@mpifr-bonn.mpg.de](mailto:reatough@mpifr-bonn.mpg.de)).

# Efficient planar heterojunction perovskite solar cells by vapour deposition

Mingzhen Liu<sup>1</sup>, Michael B. Johnston<sup>1</sup> & Henry J. Snaith<sup>1</sup>

Many different photovoltaic technologies are being developed for large-scale solar energy conversion<sup>1–4</sup>. The wafer-based first-generation photovoltaic devices<sup>1</sup> have been followed by thin-film solid semiconductor absorber layers sandwiched between two charge-selective contacts<sup>3</sup> and nanostructured (or mesostructured) solar cells that rely on a distributed heterojunction to generate charge and to transport positive and negative charges in spatially separated phases<sup>4–6</sup>. Although many materials have been used in nanostructured devices, the goal of attaining high-efficiency thin-film solar cells in such a way has yet to be achieved<sup>7</sup>. Organometal halide perovskites have recently emerged as a promising material for high-efficiency nanostructured devices<sup>8–11</sup>. Here we show that nanostructuring is not necessary to achieve high efficiencies with this material: a simple planar heterojunction solar cell incorporating vapour-deposited perovskite as the absorbing layer can have solar-to-electrical power conversion efficiencies of over 15 per cent (as measured under simulated full sunlight). This demonstrates that perovskite absorbers can function at the highest efficiencies in simplified device architectures, without the need for complex nanostructures.

Within a solar cell there are many different components with discrete roles and having different tolerances for purity and optoelectronic properties. The hybrid inorganic–organic solar cell concept is ‘material agnostic’ in that it aims to use the optimum material for each individual function. Any material that is easy to process, inexpensive and abundant can be used, with the aim of delivering a high-efficiency solar cell. Hybrid solar cells have been demonstrated in  $\pi$ -conjugated polymer blends containing semiconductor nanocrystals such as CdSe (ref. 12), CuInS<sub>2</sub> (ref. 13) and PbS (ref. 14). Dye-sensitized solar cells are hybrid solar cells containing a mesostructured inorganic n-type oxide (such as TiO<sub>2</sub>) sensitized with an organic or metal complex dye, and infiltrated with an organic p-type hole-conductor<sup>4</sup>. Recently, organometal trihalide perovskite absorbers with the general formula (RNH<sub>3</sub>)BX<sub>3</sub> (where R is C<sub>n</sub>H<sub>2n+1</sub>, X is the halogen I, Br or Cl, and B is Pb or Sn)<sup>15</sup>, have been used instead of the dye in dye-sensitized solar cells to deliver solid-state solar cells with a power conversion efficiency of over 10% (refs 8, 11, 16).

Evolving from the dye-sensitized solar cells, we found that replacing the mesoporous TiO<sub>2</sub> with mesoporous Al<sub>2</sub>O<sub>3</sub> resulted in a significant improvement in efficiency, delivering an open-circuit voltage of over 1.1 V in a device which we term a ‘meso-superstructured solar cell’<sup>8</sup>. We reason that this observed enhancement in open-circuit voltage is due to confinement of the photo-excited electrons within the perovskite phase, thereby increasing the splitting of the quasi-Fermi levels for electrons and holes under illumination, which is ultimately responsible for generating the open-circuit voltage. Further removal of the thermal sintering of the mesoporous Al<sub>2</sub>O<sub>3</sub> layer, and better optimization of processing, has led to meso-superstructured solar cells with more than 12% efficiency<sup>17</sup>. In addition, CH<sub>3</sub>NH<sub>3</sub>PbI<sub>3</sub>–<sub>x</sub>Cl<sub>x</sub> can operate relatively efficiently as a thin-film absorber in a solution-processed planar heterojunction solar cell configuration, delivering around 5% efficiency when no mesostructure is involved<sup>17</sup>. This previous work demonstrates that the perovskite absorber is capable of operating in a much simpler

planar architecture, but raises the question of whether mesostructure is essential for the highest efficiencies, or whether the thin-film planar heterojunction will lead to a superior technology.

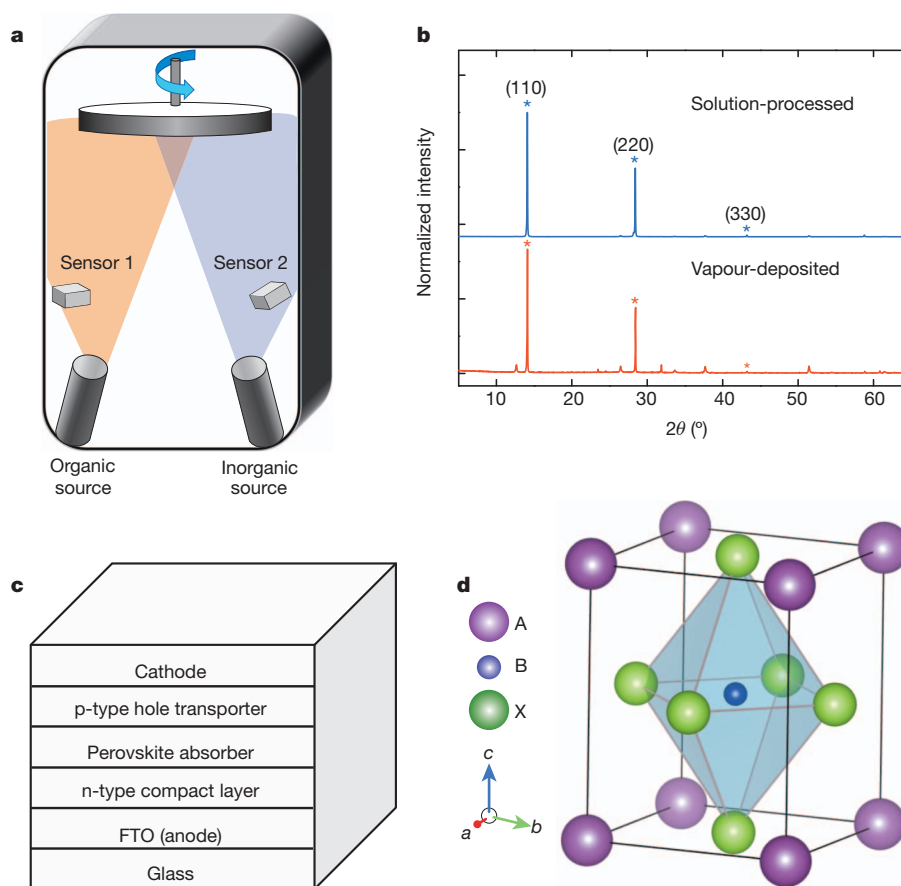
Here, as a means of creating uniform flat films of the mixed halide perovskite CH<sub>3</sub>NH<sub>3</sub>PbI<sub>3</sub>–<sub>x</sub>Cl<sub>x</sub>, we use dual-source vapour deposition. In Fig. 1 we show an illustration of the vapour-deposition set-up, along with an illustration of a planar heterojunction p–i–n solar cell (see Fig. 1c). From the bottom (the side from which the light is incident), the device is constructed on fluorine-doped tin oxide (FTO)-coated glass, coated with a compact layer of n-type TiO<sub>2</sub> that acts as the electron-selective contact. The perovskite layer is then deposited on the n-type compact layer, followed by the p-type hole conductor, 2,2',7,7'-tetrakis-(N,N-di-p-methoxyphenylamine)9,9'-spirobifluorene (spiro-OMeTAD), which ensures the selective collection of holes at the silver cathode. Given that the purpose of this study was to understand and optimize the properties of the vapour-deposited perovskite absorber layer, the compact TiO<sub>2</sub> and the spiro-OMeTAD hole transporter were solution-processed, as is usual in meso-superstructured solar cells<sup>17</sup>.

In Fig. 1b, we compare the X-ray diffraction pattern of films of CH<sub>3</sub>NH<sub>3</sub>PbI<sub>3</sub>–<sub>x</sub>Cl<sub>x</sub> either vapour-deposited or solution-cast onto compact TiO<sub>2</sub>-coated FTO-coated glass. The main diffraction peaks, assigned to the 110, 220 and 330 peaks at 14.12°, 28.44° and, respectively, 43.23°, are in identical positions for both solution-processed and vapour-deposited films, indicating that both techniques have produced the same mixed-halide perovskite with an orthorhombic crystal structure<sup>8</sup>. Notably, looking closely in the region of the (110) diffraction peak at 14.12°, there is only a small signature of a peak at 12.65° (the (001) diffraction peak for PbI<sub>2</sub>) and no measurable peak at 15.68° (the (110) diffraction peak for CH<sub>3</sub>NH<sub>3</sub>PbCl<sub>3</sub>), indicating a high level of phase purity. A diagram of the crystal structure is shown in Fig. 1d. The main difference between CH<sub>3</sub>NH<sub>3</sub>PbI<sub>3</sub> and the mixed-halide perovskite presented here is evident in a slight contraction of the *c* axis. This is consistent with the Cl atoms in the mixed-halide perovskite residing in the apical positions, out of the PbI<sub>4</sub> plane, as opposed to in the equatorial octahedral sites, as has been theoretically predicted<sup>18</sup>.

We now make a comparison between the thin-film topology and cross-sectional structure of devices fabricated by either vapour deposition or solution processing. The top-view scanning electron microscope (SEM) images in Fig. 2a, b highlight the considerable differences between the film morphologies produced by the two deposition processes. The vapour-deposited films are extremely uniform, with what appear to be crystalline features on the length scale of hundreds of nanometres. In contrast, the solution-processed films appear to coat the substrate only partially, with crystalline ‘platelets’ on the length scale of tens of micrometres. The voids between the crystals in the solution-processed films appear to extend directly to the compact TiO<sub>2</sub>-coated FTO-coated glass.

The cross-sectional images of the completed devices in Fig. 2c, d reveal more information about the crystal size. The vapour-deposited perovskite film (Fig. 2c) is uniform and similar in appearance to the FTO layer, albeit with slightly larger crystal features. The solution-processed perovskite film (Fig. 2d) is extremely smooth in the SEM

<sup>1</sup>Department of Physics, University of Oxford, Clarendon Laboratory, Parks Road, Oxford OX1 3PU, UK.



**Figure 1 | Material deposition system and characterization.** **a**, Dual-source thermal evaporation system for depositing the perovskite absorbers; the organic source was methylammonium iodide and the inorganic source  $\text{PbCl}_2$ . **b**, X-ray diffraction spectra of a solution-processed perovskite film (blue) and vapour-deposited perovskite film (red). The baseline is offset for ease of comparison

image, consistent with much larger crystal grain size than the field of view. For both of these films the crystal sizes are larger than we are able to determine from the peak width of the X-ray diffraction spectra (about 400 nm), owing to machine broadening. On zooming out, the vapour-deposited film in Fig. 2e remains flat, having an average film thickness of approximately 330 nm. In contrast, the solution-processed film in Fig. 2f has an undulating nature, with film thickness varying from 50 to 410 nm. Notably, this cross-section is still within a single 'platelet', and so even greater long-range roughness occurs owing to the areas where the perovskite absorber is completely absent (a thickness variation from 0 to 465 nm was observed in multiple SEM images).

The current-density/voltage curves measured under simulated AM1.5,  $101 \text{ mW cm}^{-2}$  irradiance (simulated sunlight) for the best-performing vapour-deposited and solution-processed planar heterojunction solar cells are shown in Fig. 3. The most efficient vapour-deposited perovskite device had a short-circuit photocurrent of  $21.5 \text{ mA cm}^{-2}$ , an open-circuit voltage of 1.07 V and a fill factor of 0.68, yielding an efficiency of 15.4%. In the same batch, the best solution-processed planar heterojunction perovskite solar cell produced a short-circuit photocurrent of  $17.6 \text{ mA cm}^{-2}$ , an open-circuit voltage of 0.84 V and a fill factor of 0.58, yielding an overall efficiency of 8.6%. In Table 1 we show the extracted performance parameters for these best-performing cells and the average with standard deviation of a batch of 12 vapour-deposited perovskite solar cells fabricated in an identical manner to the best-performing cell.

Dual-source vapour deposition results in superior uniformity of the coated perovskite films over a range of length scales, which subsequently results in substantially improved solar cell performance. In optimizing planar heterojunction perovskite solar cells, perovskite film thickness

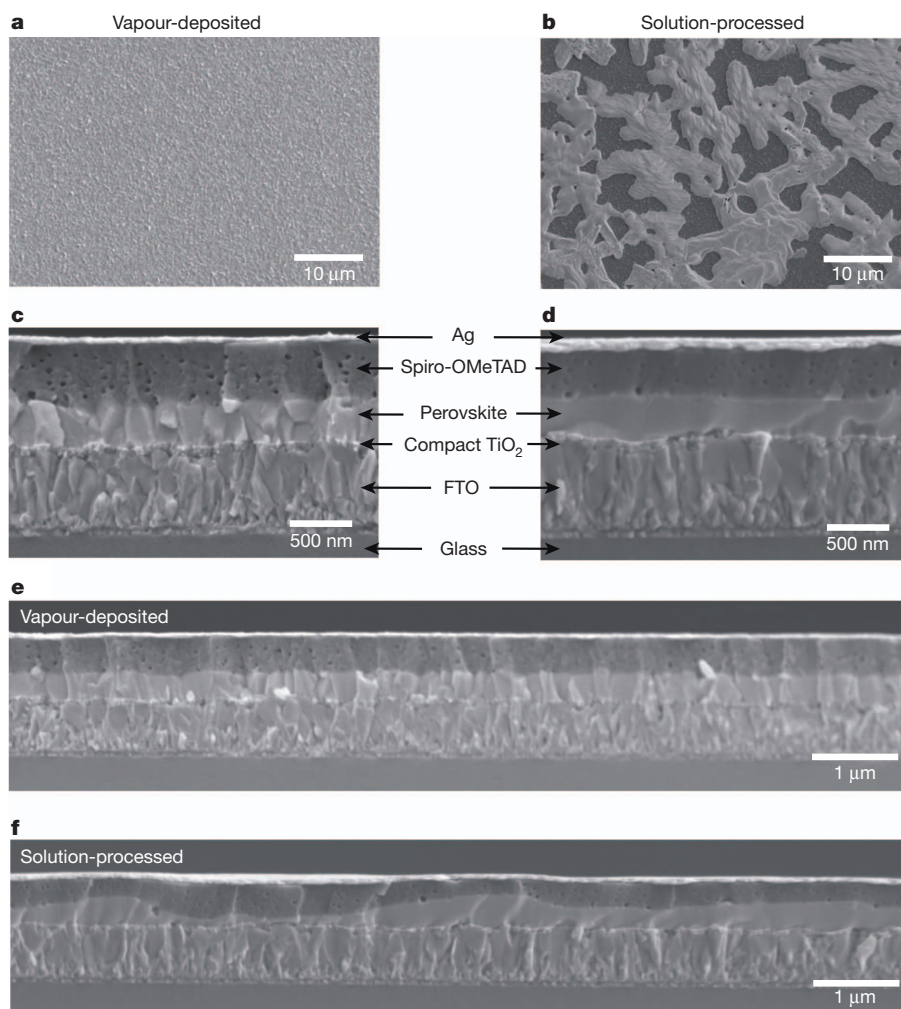
and the intensity has been normalized. **c**, Generic structure of a planar heterojunction p–i–n perovskite solar cell. **d**, Crystal structure of the perovskite absorber adopting the perovskite  $\text{ABX}_3$  form, where A is methylammonium, B is Pb and X is I or Cl.

is a key parameter. If the film is too thin, then that region will not absorb sufficient sunlight. If the film is too thick, there is a significant chance that the electron and hole (or exciton) diffusion length will be shorter than the film thickness, and that the charge will therefore not be collected at the p-type and n-type heterojunctions. Furthermore, the complete absence of material from some regions in the solution-processed films (pinholes) will result in direct contact of the p-type spiro-OMeTAD and the  $\text{TiO}_2$  compact layer. This leads to a shunting path that is probably partially responsible for the lower fill factor and open-circuit voltage in the solution-cast planar heterojunction devices<sup>17,19</sup>. Indeed, it is remarkable that such inhomogeneous and undulating solution-cast films can deliver devices with over 8% efficiency.

The results presented here demonstrate that solid perovskite layers can operate extremely well in a solar cell, and in essence set a lower limit of 330 nm (the film thickness) on the electron and hole diffusion length in this perovskite absorber. However, more work is required to determine the electron and hole diffusion lengths precisely and to understand the primary excitation and the mechanisms for free-charge generation in these materials.

A distinct advantage of vapour deposition over solution processing is the ability to prepare layered multi-stack thin films over large areas. Vapour deposition is a mature technique used in the glazing industry, the liquid-crystal display industry and the thin-film solar cell industry, among others. Vapour deposition can lead to full optimization of electronic contact at interfaces through multilayers with controlled levels of doping<sup>20</sup>, as is done in the crystalline silicon 'heterojunction with thin intrinsic layer' solar cell<sup>21</sup> and in thin-film solar cells<sup>3</sup>. Additionally, organic light-emitting diodes<sup>22,23</sup> have proved to be commercially sound, with devices with extremely thin multilayer stacks fabricated by vapour

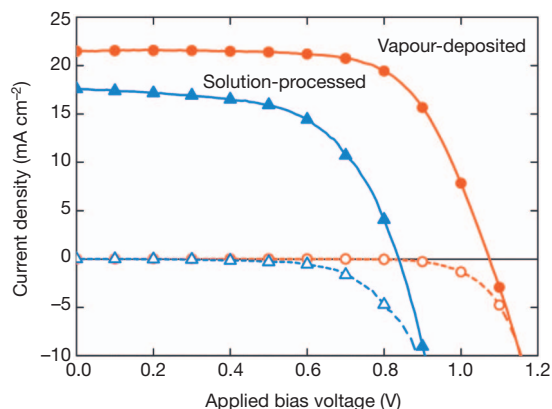




**Figure 2 | Thin-film topology characterization.** **a, b**, SEM top views of a vapour-deposited perovskite film (**a**) and a solution-processed perovskite film (**b**). **c, d**, Cross-sectional SEM images under high magnification of complete solar cells constructed from a vapour-deposited perovskite film (**c**) and a

solution-processed perovskite film (**d**). **e, f**, Cross-sectional SEM images under lower magnification of completed solar cells constructed from a vapour-deposited perovskite film (**e**) and a solution-processed perovskite film (**f**).

deposition. Small molecular organic photovoltaics have also been able to compete directly with solution-processed organic photovoltaics despite much lower levels of research and development, because with



**Figure 3 | Solar cell performance.** Current-density/voltage curves of the best-performing solution-processed (blue lines, triangles) and vapour-deposited (red lines, circles) planar heterojunction perovskite solar cells measured under simulated AM1.5 sunlight of  $101 \text{ mW cm}^{-2}$  irradiance (solid lines) and in the dark (dashed lines). The curves are for the best-performing cells measured and their reproducibility is shown in Table 1.

vapour deposition the charge-collection interfaces can be carefully tuned, and multi-junction architectures are more straightforward to realize<sup>24</sup>. An interesting possibility for the current vapour-deposited perovskite technology is to use it as a ‘top cell’ in a hybrid tandem junction with either crystalline silicon or copper indium gallium (di)selenide. Although ultimately an ‘all-perovskite’ multi-junction cell should be realizable, the perovskite cells have now achieved a performance that is sufficient to increase the absolute efficiency of high-efficiency crystalline silicon and copper indium gallium (di)selenide solar cells<sup>25</sup>. Additionally, because vapour deposition of the perovskite layers is entirely compatible with conventional processing methods for silicon-wafer-based and thin-film solar cells, the infrastructure could already be in place to scale up this technology.

We have built vapour-deposited organometal trihalide perovskite solar cells based on a planar heterojunction thin-film architecture that have a solar-to-electrical power conversion efficiency of over 15% with an open-circuit voltage of 1.07 V. The perovskite absorbers seem to be versatile materials for incorporation into highly efficient solar cells, given the low-temperature processing they require, the option of using either solution processing or vapour deposition or both, the simplified device architecture and the availability of many other metal and organic salts that could form a perovskite structure. Whether vapour deposition emerges as the preferred route for manufacture or simply represents a benchmark method for fabricating extremely uniform

**Table 1 | Solar cell performance parameters**

	Current density (mA cm <sup>-2</sup> )	Open-circuit voltage (V)	Fill factor	Efficiency (%)
Vapour-deposited	21.5	1.07	0.67	15.4
Vapour-deposited (average ± s.d.)	18.9 ± 1.8	1.05 ± 0.03	0.62 ± 0.05	12.3 ± 2.0
Solution-processed	17.6	0.84	0.58	8.6

Solar cell performance parameters are extracted from the current–voltage curves presented in Fig. 3. The average and standard deviation (s.d) values of a batch of 12 identically processed planar heterojunction vapour-deposited CH<sub>3</sub>NH<sub>3</sub>PbI<sub>3-x</sub>Cl<sub>x</sub> perovskite solar cells are also shown.

films (that will ultimately be matched by solution processing) remains to be seen. Finally, a key target for the photovoltaics community has been to find a wider-bandgap highly efficient ‘top cell’, to enable the next step in improving the performance of crystalline silicon and existing second-generation thin-film solar cells. This perovskite technology is now compatible with these first- and second-generation technologies, and is hence likely to be adopted by the conventional photovoltaics community and industry. Therefore, it may find its way rapidly into utility-scale power generation.

## METHODS SUMMARY

The perovskite absorber was deposited by a dual-source evaporation system (Kurt J. Lesker Mini Spectros) with ceramic crucibles (organic light-emitting diode sources) in a nitrogen-filled glovebox. The vapour-deposited perovskite devices were fabricated on FTO-coated glass. A compact layer of TiO<sub>2</sub> was deposited on the FTO-coated glass by spin-coating (solution-processing) it with a mildly acidic solution of titanium isopropoxide in ethanol<sup>17</sup>, and subsequently the perovskite absorber was deposited on the compact TiO<sub>2</sub>-coated FTO substrate. Methylammonium iodide (CH<sub>3</sub>NH<sub>3</sub>I) and lead chloride (PbCl<sub>2</sub>) were the organic and inorganic precursor salts, evaporated simultaneously from separate sources at 10<sup>-5</sup> mbar with an as-deposited molar ratio of 4:1, based on the reading of the sensors above the crucibles. A dark reddish-brown colour was observed immediately after evaporation. Annealing the perovskite absorbers before spin-coating the hole-transporter layer fully crystallized the perovskite layer. After spin-coating the hole transporter, spiro-OMeTAD, from a chlorobenzene solution to form the photoactive layer (including lithium bis(trifluoromethylsulfonyl)imide salt and tert-butylpyridine as additives)<sup>8</sup>, the devices were capped with silver metal electrodes through thermal evaporation at 10<sup>-6</sup> mbar. Full details of material and device fabrication and characterization techniques are available in Methods. All other characterizations and measurements were carried out as previously described<sup>17</sup>.

**Online Content** Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 19 June; accepted 25 July 2013.**

**Published online 11 September 2013.**

- Green, M. A. Silicon photovoltaic modules: a brief history of the first 50 years. *Prog. Photovolt. Res. Appl.* **13**, 447–455 (2005).
- Graetzel, M., Janssen, R. A. J., Mitzi, D. B. & Sargent, E. H. Materials interface engineering for solution-processed photovoltaics. *Nature* **488**, 304–312 (2012).
- Chopra, K. L., Paulson, P. D. & Dutta, V. Thin-film solar cells: an overview. *Prog. Photovolt. Res. Appl.* **12**, 69–92 (2004).
- O'Regan, B. & Grätzel, M. A low-cost, high-efficiency solar cell based on dye-sensitized colloidal TiO<sub>2</sub> films. *Nature* **353**, 737–740 (1991).
- Halls, J. J. M. *et al.* Efficient photodiodes from interpenetrating polymer networks. *Nature* **376**, 498–500 (1995).
- Yu, G., Gao, J., Hummelen, J. C., Wudl, F. & Heeger, A. J. Polymer photovoltaic cells: enhanced efficiencies via a network of internal donor-acceptor heterojunctions. *Science* **270**, 1789–1791 (1995).
- Green, M. A., Emery, K., Hishikawa, Y., Warta, W. & Dunlop, E. D. Solar cell efficiency tables (version 40). *Prog. Photovolt. Res. Appl.* **20**, 606–614 (2012).

- Lee, M. M., Teuscher, J., Miyasaka, T., Murakami, T. N. & Snaith, H. J. Efficient hybrid solar cells based on meso-superstructured organometal halide perovskites. *Science* **338**, 643–647 (2012).
- Kojima, A., Teshima, K., Shirai, Y. & Miyasaka, T. Organometal halide perovskites as visible-light sensitizers for photovoltaic cells. *J. Am. Chem. Soc.* **131**, 6050–6051 (2009).
- Kim, H.-S. *et al.* Lead iodide perovskite sensitized all-solid-state submicron thin film mesoscopic solar cell with efficiency exceeding 9%. *Sci. Rep.* **2**, 591, doi:10.1038/srep00591 (2012).
- Noh, J. H., Im, S. H., Heo, J. H., Mandal, T. N. & Il Seok, S. Chemical management for colorful, efficient, and stable inorganic-organic hybrid nanostructured solar cells. *Nano Lett.* **13**, 1764–1769 (2013).
- Huynh, W. U., Dittmer, J. J. & Alivisatos, A. P. Hybrid nanorod-polymer solar cells. *Science* **295**, 2425–2427 (2002).
- Arci, E., Sariciftci, N. S. & Meissner, D. Hybrid solar cells based on nanoparticles of CuInS<sub>2</sub> in organic matrices. *Adv. Funct. Mater.* **13**, 165–171 (2003).
- McDonald, S. A. *et al.* Solution-processed PbS quantum dot infrared photodetectors and photovoltaics. *Nature Mater.* **4**, 138–142 (2005).
- Cheng, Z. & Lin, J. Layered organic–inorganic hybrid perovskites: structure, optical properties, film preparation, patterning and templating engineering. *CrystEngComm* **12**, 2646–2662 (2010).
- Heo, J. H. *et al.* Efficient inorganic-organic hybrid heterojunction solar cells containing perovskite compound and polymeric hole conductors. *Nature Photon.* **7**, 486–491 (2013).
- Ball, J. M., Lee, M. M., Hey, A. & Snaith, H. J. Low-temperature processed meso-superstructured to thin-film perovskite solar cells. *Energy Environ. Sci.* **6**, 1739–1743 (2013).
- Mosconi, E., Amat, A., Nazeeruddin, M. K., Grätzel, M. & De Angelis, F. First principles modeling of mixed halide organometal perovskites for photovoltaic applications. *J. Phys. Chem. C* **117**, 13902–13913 (2013).
- Snaith, H. J., Greenham, N. C. & Friend, R. H. The origin of collected charge and open-circuit voltage in blended polyfluorene photovoltaic devices. *Adv. Mater.* **16**, 1640–1645 (2004).
- Yang, F., Shtein, M. & Forrest, S. R. Morphology control and material mixing by high-temperature organic vapor-phase deposition and its application to thin-film solar cells. *J. Appl. Phys.* **98**, 014906 (2005).
- Sakata, H. *et al.* in *Photovoltaic Specialists Conf. (Conference Record of the Twenty-Eighth IEEE)* 7–12, doi:10.1109/PVSC.2000.915742 (2000).
- Reineke, S. *et al.* White organic light-emitting diodes with fluorescent tube efficiency. *Nature* **459**, 234–238 (2009).
- You, H., Dai, Y., Zhang, Z. & Ma, D. Improved performances of organic light-emitting diodes with metal oxide as anode buffer. *J. Appl. Phys.* **101**, 026105 (2007).
- Riede, M. *et al.* Efficient organic tandem solar cells based on small molecules. *Adv. Funct. Mater.* **21**, 3019–3028 (2011).
- Beiley, Z. M. & McGehee, M. D. Modeling low cost hybrid tandem photovoltaics with the potential for efficiencies exceeding 20%. *Energy Environ. Sci.* **5**, 9173–9179 (2012).

**Acknowledgements** This work was funded by EPSRC and the European Research Council (ERC) ‘Hyper Project’ number 279881. The Oxford University Press (John Fell) Fund provided support for equipment used in this study, specifically the organic light-emitting diode vapour-deposition equipment. We thank S. Sun, E. Crossland, P. Docampo, G. Eperon, J. Zhang and J. Liu for discussions and experimental and technical assistance.

**Author Contributions** M.L. performed the experimental work, data analysis and experimental planning. The project was conceived, planned and supervised by H.S. and M.J. The manuscript was written by all three authors.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to H.J.S. ([h.snaith1@physics.ox.ac.uk](mailto:h.snaith1@physics.ox.ac.uk)).

## METHODS

**Device substrate preparation.** Substrate preparation was undertaken under ambient conditions. TEC7 Glass FTO-coated glass (TEC7,  $7\ \Omega/\square$  sheet resistivity) was patterned by etching with Zn metal powder and 2 M HCl diluted in deionized water. The substrates were then cleaned with a 2% solution of Hellmanex cuvette cleaning detergent diluted in deionized water, rinsed with deionized water, acetone and ethanol, and dried with clean dry air<sup>8</sup>. Oxygen plasma was subsequently used to treat the substrate for 10 min. An acidic solution of titanium isopropoxide in ethanol was spin-coated (solution-processed) onto the clean substrates at 2,000 r.p.m. for 1 min, before drying at 150 °C for 10 min and then sintering at 500 °C for 30 min to form a compact n-type layer of TiO<sub>2</sub> (ref. 17).

**Vapour deposition of perovskite absorber from precursor salts.** Subsequently, the perovskite absorbers were deposited through dual-source evaporation from lead chloride (PbCl<sub>2</sub>) and methylammonium iodide (CH<sub>3</sub>NH<sub>3</sub>I) simultaneously, onto the TiO<sub>2</sub>-compact-layer-coated FTO substrates under high vacuum.

**Tooling factor estimation.** Each vapour-deposition source was monitored using a quartz crystal monitor positioned a short distance from the source. The source-to-monitor distance is different to the source-to-substrate distance, and so a tooling factor (which is a ratio of the material deposited on the sensors to that on the samples) was estimated for each source individually before dual-source evaporation. The tooling factor for both sources was initially set to 1 on the evaporator programme. The parameters required for estimating the thickness of the deposited material are shown in Extended Data Table 1. We note that the density and acoustic impedance of the CH<sub>3</sub>NH<sub>3</sub>I were unknown but both set to be 1 (the density of CH<sub>3</sub>NH<sub>3</sub>Cl has been reported as being 1.1 g cm<sup>-3</sup> so this is probably close, and any absolute difference is accommodated by using the tooling factor). Initially, we ramped the temperature of each source individually, and simultaneously measured the deposition rate to gain an appreciation of the temperature range and range of deposition rates feasible for each material. We then chose a temperature for each material which would give a reasonable thickness of deposited material over a 30-min deposition period. We then carried out an evaporation of each source for a total time of approximately 30 min at a fixed evaporation rate, with the final deposited thickness recorded on the crystal monitor. The actual deposited thickness on the substrate we then measured with a surface profilometer and the tooling factor was estimated by dividing the sensor-estimated thickness by the measured deposited thickness on the substrate. The tooling results are shown in Extended Data Table 1. These tooling factors were then subsequently applied during co-depositions. The deposition rate for CH<sub>3</sub>NH<sub>3</sub>I typically varied by  $\pm 15\%$  during an evaporation, whereas the rate for PbCl<sub>2</sub> varied by  $\pm 10\%$ .

**Dual-source evaporation.** We placed approximately 500 mg of CH<sub>3</sub>NH<sub>3</sub>I and 100 mg of PbCl<sub>2</sub> into separate crucibles. The device substrates were placed in a substrate holder above the sources with the TiO<sub>2</sub>-coated FTO side facing down towards the sources. Once the pressure in the chamber was pumped down to below  $10^{-5}$  mbar, the two sources were heated slightly above their desired deposition temperatures for approximately 5 min (that is, CH<sub>3</sub>NH<sub>3</sub>I was heated to about 120 °C and PbCl<sub>2</sub> was heated to about 325 °C) to remove volatile impurities before depositing the materials onto the substrate. The substrate holder was rotated to ensure uniform coating throughout deposition, because the right-hand source predominantly coats the right-hand side of the substrate and similarly for the left. The substrate holder was water-cooled to approximately 21 °C, though precise measurement of the substrate temperature during deposition was not performed. Perovskite films were optimized for best device performance by varying the key deposition parameters such as the deposition rates and times for the

two sources. In particular, device performance was very sensitive to the relative composition of CH<sub>3</sub>NH<sub>3</sub>I to PbCl<sub>2</sub> as well as to the overall deposited thickness.

In our trials we performed the following steps. (1) We varied the as-deposited composition of CH<sub>3</sub>NH<sub>3</sub>I to PbCl<sub>2</sub> from 1:1 to 7:1, at a fixed as-annealed film thickness of 125 nm. (2) We varied the film thickness—at the optimum ratio of CH<sub>3</sub>NH<sub>3</sub>I:PbCl<sub>2</sub> = 3.5:1—from 125 to 500 nm and found the optimum performance at 330 nm. (3) We fine-tuned the CH<sub>3</sub>NH<sub>3</sub>I:PbCl<sub>2</sub> ratio for films with about 330 nm thickness to obtain an optimum composition of 4:1 CH<sub>3</sub>NH<sub>3</sub>I:PbCl<sub>2</sub>. (4) We optimized the hole-transporter thickness (solution concentration) and Li-TFSI dopant concentration to maximize performance on the  $\sim 330$ -nm-thick, 4:1 CH<sub>3</sub>NH<sub>3</sub>I:PbCl<sub>2</sub> deposited perovskite films. The optimal deposition rate was  $5.3\ \text{\AA s}^{-1}$  for CH<sub>3</sub>NH<sub>3</sub>I (achieved with a crucible temperature of around 116 °C) and  $1\ \text{\AA s}^{-1}$  for PbCl<sub>2</sub> (achieved with a crucible temperature of around 320 °C), maintained for approximately 128 min of evaporation as shown in Extended Data Table 2.

The colour of the samples after deposition varied depending on the composition of the two sources. For the best-performing devices, a reddish-brown colour was observed and the film appears to be partially crystallized in the topological SEM image of the as-deposited film, as shown in Extended Data Fig. 1. Annealing the as-deposited films at 100 °C for 45 min in the N<sub>2</sub>-filled glove box before spin-coating the hole transporter enabled full crystallization of the perovskite, darkening the colour and resulting in an apparent growth of the crystal features visible in the SEM image, as shown in Extended Data Fig. 1. After annealing, the best-performing samples had an average thickness of approximately 330 nm. As would be expected from the non-stoichiometric as-deposited molar ratio, there is clearly significant mass loss and hence thickness loss during perovskite film formation.

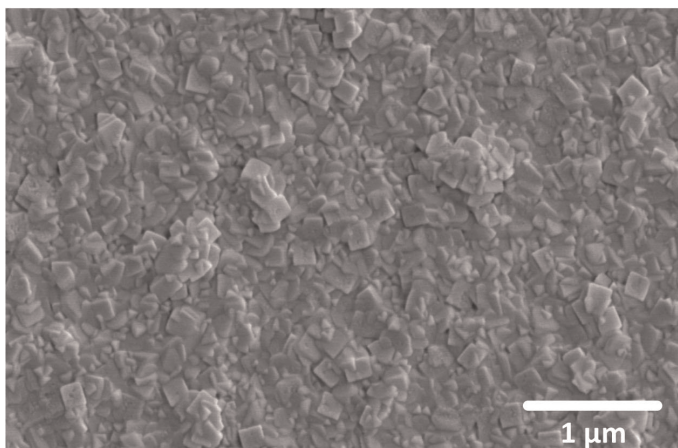
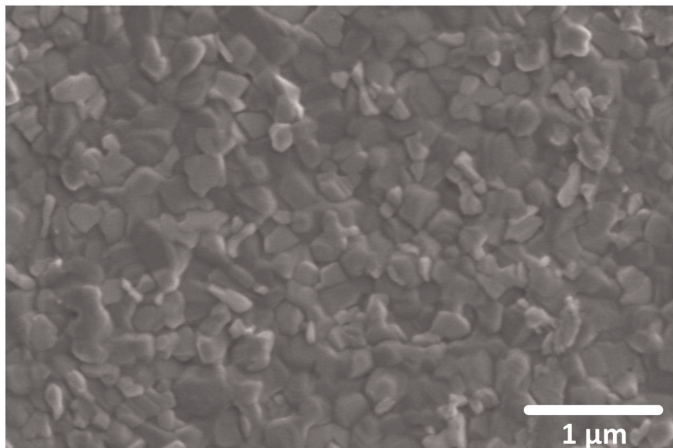
**Hole-transporter deposition.** The hole-transporter layer was deposited by spin-coating (2,000 r.p.m. for 45 s) 25  $\mu\text{l}$  of chlorobenzene solution that contained 61.4 mM spiro-OMeTAD, 55 mM tert-butylpyridine (tBP) and 26 mM lithium bis(trifluoromethylsulfonyl)imide salt. Before completing the device fabrication by the thermal evaporation of a silver cathode, the devices were left in a desiccator overnight and the unsealed devices were tested in air immediately after the cathode fabrication. Note that the vapour deposition of the perovskite absorbers and the spin coating of the hole transporter layer were both done in the N<sub>2</sub>-filled glove box.

**X-ray diffraction.**  $2\theta$  scans were obtained from samples of perovskite deposited on the compact-TiO<sub>2</sub>-coated FTO-coated glass using an X-ray diffractometer (Panalytical X'Pert Pro).

**SEM images.** An emission SEM (Hitachi S-4300 field) was used for collecting the SEM images. The instrument uses an electron beam accelerated at 500 V to 30 kV, enabling operation at a variety of currents.

**Current-voltage characteristics.** Current-voltage characteristics were measured (2400 Series SourceMeter, Keithley Instruments) under simulated AM1.5 sunlight at  $101\ \text{mW cm}^{-2}$  irradiance generated by an Abet Class AAB Sun 2000 simulator, with the intensity calibrated with a National Renewables Energy Laboratory calibrated KG5-filtered Si reference cell (the rated error in the short-circuit photocurrent on the reference cell is 1.36% to 95% certainty). The mismatch factor was estimated to be less than 1%. The solar devices were masked with a metal aperture to define the active area of about  $0.076 \pm 0.002\ \text{cm}^2$  (error in mask area estimated from multiple measurements of a number of different masks, designed to be the same size) and measured in a light-tight sample holder to minimize any edge effects and ensure that both the reference and test cells were located in the same spot under the solar simulator during measurement.



**a****b**

**Extended Data Figure 1** | Top-view SEM images for the vapour-deposited perovskite films. **a**, As-deposited perovskite film; **b**, post-annealed perovskite film.

**Extended Data Table 1 | Tooling factor measurement of the dual-source vapour-deposition system.**

Source	Density	Evaporation rate	Values on sensors	Actual film thickness (mean)	Tooling factor
Source 1: CH <sub>3</sub> NH <sub>3</sub> I	1 g cm <sup>-3</sup> *	15 Ås <sup>-1</sup>	30 kÅ	1.39 µm	2.16
Source 2: PbCl <sub>2</sub>	5.85 g cm <sup>-3</sup>	5 Ås <sup>-1</sup>	10 kÅ	185 nm	5.41

\* The density of CH<sub>3</sub>NH<sub>3</sub>I is assumed to be 1 g cm<sup>-3</sup> because its precise density is not known.

Extended Data Table 2 | Optimized deposition conditions for the evaporated perovskite solar devices.

Source	Evaporation rate	Values on sensors	Tooling factor	Actual deposited equivalent thickness	Approximate molar ratio on substrate (CH <sub>3</sub> NH <sub>3</sub> I : PbCl <sub>2</sub> )
Source 1: CH <sub>3</sub> NH <sub>3</sub> I	5.3±0.8 Ås <sup>-1</sup>	42.4 kÅ	2.16	1.96 μm	4:1
Source 2: PbCl <sub>2</sub>	1 Ås <sup>-1</sup>	8 kÅ	5.41	148 nm	

We have used the densities shown in Extended Data Table 1 and the relative molecular mass of 157 for CH<sub>3</sub>NH<sub>3</sub>I and 278 for PbCl<sub>2</sub> to calculate the relative molar ratio of as-deposited material.



# Ultrastable silver nanoparticles

Anil Desireddy<sup>1</sup>, Brian E. Conn<sup>1</sup>, Jingshu Guo<sup>1</sup>, Bokwon Yoon<sup>2</sup>, Robert N. Barnett<sup>2</sup>, Bradley M. Monahan<sup>1</sup>, Kristin Kirschbaum<sup>1</sup>, Wendell P. Griffith<sup>1</sup>, Robert L. Whetten<sup>3,4</sup>, Uzi Landman<sup>2</sup> & Terry P. Bigioni<sup>1,5</sup>

Noble-metal nanoparticles have had a substantial impact across a diverse range of fields, including catalysis<sup>1</sup>, sensing<sup>2</sup>, photochemistry<sup>3</sup>, optoelectronics<sup>4,5</sup>, energy conversion<sup>6</sup> and medicine<sup>7</sup>. Although silver has very desirable physical properties, good relative abundance and low cost, gold nanoparticles have been widely favoured owing to their proved stability and ease of use. Unlike gold, silver is notorious for its susceptibility to oxidation (tarnishing), which has limited the development of important silver-based nanomaterials. Despite two decades of synthetic efforts, silver nanoparticles that are inert or have long-term stability remain unrealized. Here we report a simple synthetic protocol for producing ultrastable silver nanoparticles, yielding a single-sized molecular product in very large quantities with quantitative yield and without the need for size sorting. The stability, purity and yield are substantially better than those for other metal nanoparticles, including gold, owing to an effective stabilization mechanism. The particular size and stoichiometry of the product were found to be insensitive to variations in synthesis parameters. The chemical stability and structural, electronic and optical properties can be understood using first-principles electronic structure theory based on an experimental single-crystal X-ray structure. Although several structures have been determined for protected gold nanoclusters<sup>8–12</sup>, none has been reported so far for silver nanoparticles. The total structure of a thiolate-protected silver nanocluster reported here uncovers the unique structure of the silver thiolate protecting layer, consisting of Ag<sub>4</sub>S<sub>5</sub> capping structures. The outstanding stability of the nanoparticle is attributed to a closed-shell 18-electron configuration with a large energy gap between the highest occupied molecular orbital and the lowest unoccupied molecular orbital, an ultrastable 32-silver-atom excavated-dodecahedral<sup>13</sup> core consisting of a hollow 12-silver-atom icosahedron encapsulated by a 20-silver-atom dodecahedron, and the choice of protective coordinating ligands. The straightforward synthesis of large quantities of pure molecular product promises to make this class of materials widely available for further research and technology development<sup>14–18</sup>.

Silver thiolate clusters with an aromatic ligand shell and a core diameter of ~1.2 nm have been discovered only recently<sup>19</sup>. Electrospray-ionization mass spectrometry (ESI-MS) identified these as discrete molecular complexes, Ag<sub>44</sub>(SPh)<sub>30</sub><sup>4–</sup>, where SPh represents aryl thiolates<sup>20</sup>. Complexity of preparation and handling have proven limiting, however, as these clusters shared the typical vulnerabilities of Ag nanoparticles.

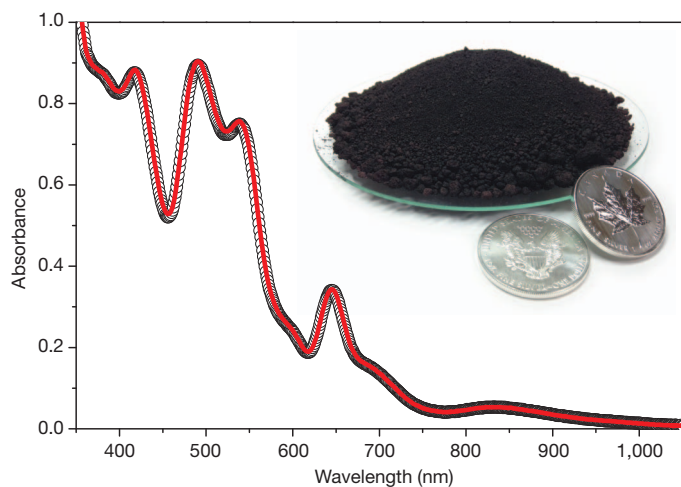
We have developed a new approach to the preparation of ultrastable silver nanoparticles in semiaqueous solution, which uses a protecting ligand shell of *p*-mercaptobenzoic acid (*p*-MBA)<sup>8,19</sup>. With judicious choice of solvent conditions and stabilizing agents, the fragile and unstable Ag complexes can be transformed into chemically inert materials with unprecedented stability. The synthesis involves the reduction of a soluble precursor in semiaqueous solution in the presence of alkali metal cations and a coordinating solvent. The straightforward protocol produces a pure molecular material without size separations and achieves near quantitative yield in large quantities (Fig. 1 inset). The product

can be dried and fully redispersed in protic, aprotic and nonpolar solvents with no loss of material or change in chemical identity.

The absorption spectrum of the raw product is highly structured (Fig. 1) and identical to that of the purified material, with an onset at about 1,100 nm (~1.1 eV). ESI-MS of the raw product (Fig. 2) identified several ion species that were all attributed to a single cluster size, the pure Ag<sub>44</sub>(*p*-MBA)<sub>30</sub><sup>4–</sup> complex (mass-to-charge ratio *m/z* 2,336). The Ag<sub>43</sub>(*p*-MBA)<sub>28</sub><sup>3–</sup> complex (*m/z* 2,975) was attributed to electrostatic destabilization and spontaneous fragmentation of Ag<sub>44</sub>(*p*-MBA)<sub>30</sub><sup>4–</sup> on desolvation (see Supplementary Information). The experimental data only matched the simulated isotopic distribution for the fully protonated species (see Supplementary Information), therefore the entire 4– charge was carried by the silver core rather than by the carboxylates. Four alkali counterions (M) were identified by elemental analysis, giving M<sub>4</sub>Ag<sub>44</sub>(*p*-MBA)<sub>30</sub> as the molecular formula.

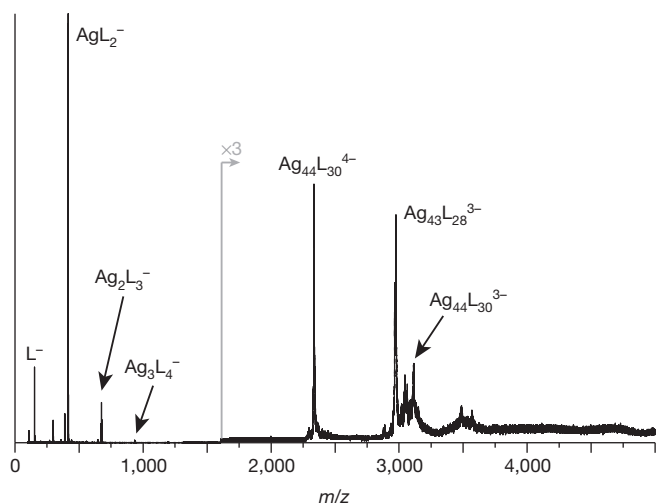
The synthesis of M<sub>4</sub>Ag<sub>44</sub>(*p*-MBA)<sub>30</sub> has some characteristics not found in other nanoparticle preparations. Single-sized products are usually isolated by attrition: the less stable sizes are either destroyed or converted into the most stable size<sup>21–23</sup>. Direct synthesis of a truly single-sized molecular product with yields >95% indicates that these clusters are more stable than any other known cluster species. Furthermore, the particular size, composition and stoichiometry of the nanocluster product was found to be immune to changes in experimental parameters (for example, solvent composition, reactant concentrations) for the synthesis method reported here.

The profound difference between canonical nanoparticle syntheses and the present work was demonstrated by synthesizing M<sub>4</sub>Ag<sub>44</sub>(*p*-MBA)<sub>30</sub>



**Figure 1 | Optical absorption and material sample.** Absorption spectrum of the M<sub>4</sub>Ag<sub>44</sub>(*p*-MBA)<sub>30</sub> raw product solution (red line) synthesized in the presence of 'seed' M<sub>4</sub>Ag<sub>44</sub>(*p*-MBA)<sub>30</sub> clusters (spectrum shown as open circles). Inset, 140 g of M<sub>4</sub>Ag<sub>44</sub>(*p*-MBA)<sub>30</sub> clusters pictured with two one-ounce silver coins for scale (each coin is 4 cm in diameter and weighs 31.1 g). The dish is 18 cm in diameter.

<sup>1</sup>Department of Chemistry, University of Toledo, Toledo, Ohio 43606, USA. <sup>2</sup>School of Physics, Georgia Institute of Technology, Atlanta, Georgia 30332-0430, USA. <sup>3</sup>School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, Georgia 30332-0400, USA. <sup>4</sup>Department of Physics and Astronomy, University of Texas at San Antonio, San Antonio, Texas 78249, USA. <sup>5</sup>School of Solar and Advanced Renewable Energy, University of Toledo, Toledo, Ohio 43606, USA.



**Figure 2 | Electrospray-ionization mass spectrum.** ESI-MS of the final product without size separation shows that only one species is present. Peaks from 3,000–3,200  $m/z$  are fragments with 3<sup>−</sup> charge state, and the broad intensity at 3,500  $m/z$  is attributed to non-specific dimerization of fragments with 5<sup>−</sup> total charge state (see Supplementary Information for further details). Here, L is *p*-MBA.

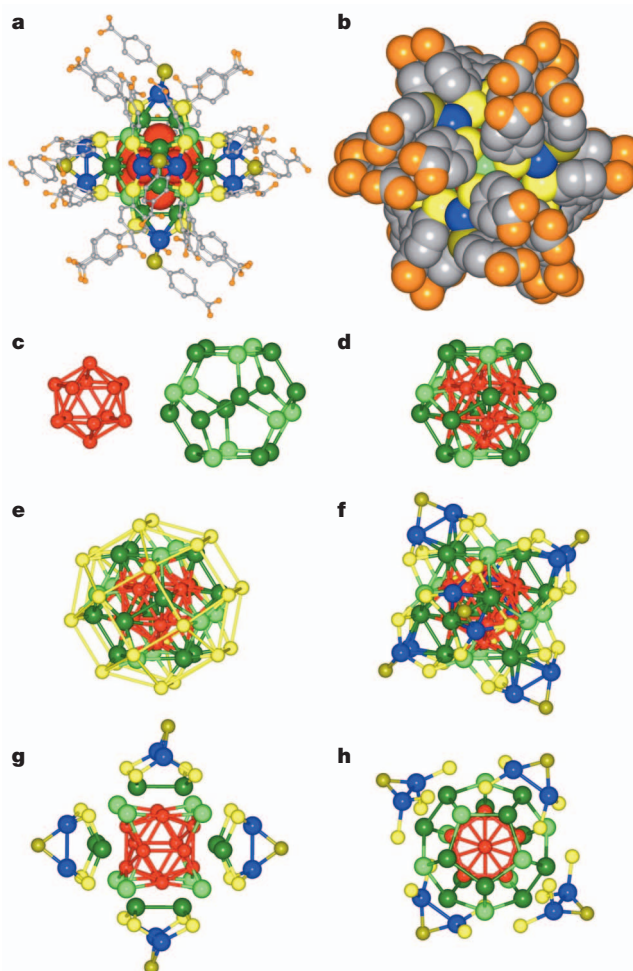
clusters in the presence of existing  $M_4Ag_{44}(p\text{-MBA})_{30}$  clusters. Normally the existing nanoparticles would act as seeds and grow at the expense of new particle nucleation<sup>24</sup>. Instead,  $M_4Ag_{44}(p\text{-MBA})_{30}$  clusters were formed with identical yield and chemical identity, with or without these seeds present (Fig. 1). Once formed,  $M_4Ag_{44}(p\text{-MBA})_{30}$  clusters were very stable and unreactive, behaving as inert molecules rather than as typical nanoparticles. When an analogous reaction was performed with  $Au_{25}(SG)_{18}$  (SG = glutathionate), the clusters were not inert but rather acted as canonical seeds (see Supplementary Information).

The long-term stability of solutions of  $M_4Ag_{44}(p\text{-MBA})_{30}$  clusters were also superior to those of  $Au_{25}(SG)_{18}$  clusters. The ambient decay rates of  $M_4Ag_{44}(p\text{-MBA})_{30}$  cluster solutions were  $\sim 7$  times slower than those of  $Au_{25}(SG)_{18}$  cluster solutions (see details in Supplementary Information and Supplementary Fig. 4b). Therefore, under both ambient (mildly oxidizing) and reducing conditions, the  $M_4Ag_{44}(p\text{-MBA})_{30}$  clusters proved to be more noble than even the highly stable  $Au_{25}(SG)_{18}$  cluster<sup>21</sup>. When discussing relative stabilities of protected nanoclusters, caution should be exercised with regard to the ligands used and the environmental conditions (see Supplementary Information). Indeed, experiments in our laboratory have shown that  $Au_{25}(p\text{-MBA})_{18}$  clusters are too unstable for meaningful temporal stability measurements to be made.

This inertness under reducing conditions implies that the synthesis of new clusters can carry on without regard for existing clusters in the reaction vessel, giving impetus to scaling up the reaction. Indeed, it has been possible to produce 140 g of the final  $M_4Ag_{44}(p\text{-MBA})_{30}$  product from a single reaction (Fig. 1 inset), although kilogram-scale syntheses should be easily achievable. We note that 140 g is three orders of magnitude larger than typical nanoparticle preparations.

Clues as to the origins of the stability of  $M_4Ag_{44}(p\text{-MBA})_{30}$  as well as the structure of the thiol surface layers that protect the silver nanoparticles have been revealed by single-crystal X-ray diffraction.  $Na_4Ag_{44}(p\text{-MBA})_{30}$  clusters were crystallized from dimethylformamide (DMF) solution, with rhombus-shaped crystals forming after 1–3 days. The entire structure of the cluster was determined by single-crystal X-ray crystallography (see Supplementary Information), and is shown in Fig. 3a, b.

The crystal structure has exceptionally high symmetry, containing elements that exhibit four of the five Platonic solids. The all-silver core consists of a hollow icosahedron ( $Ag_{12}$  inner core) within a dodecahedron ( $Ag_{20}$  outer core), forming an  $Ag_{32}$  excavated-dodecahedral



**Figure 3 | X-ray crystal structure obtained from a  $Na_4Ag_{44}(p\text{-MBA})_{30}$  crystal.** **a**, Complete cluster structure showing silver core and *p*-MBA ligands (see colour scheme below). **b**, Space-filling view down a three-fold axis (see colour scheme below). Note face-to-face and edge-to-face  $\pi$  stacking in the groupings of two and three ligands, resulting in considerable void space. **c**, The  $Ag_{32}$  excavated-dodecahedral core consists of an inner 12-atom (hollow) icosahedron (red) whose atoms do not contact sulphur, encapsulated by a 20-atom dodecahedron (green). **d**, The complete  $Ag_{32}$  excavated-dodecahedral core. The eight atoms of the dodecahedron that are coloured light green define a cube, with pairs of dark-green Ag atoms located above the faces. **e**, Sulphur atoms (yellow) are arranged in a slightly distorted rhombicuboctahedron with S atoms in the triangular faces coordinating to the light-green Ag atoms of the 20-atom dodecahedron. **f**, Six faces of the rhombicuboctahedron are capped with an  $Ag_2S$  unit with the bridging S atom tilted off axis, completing the inorganic structure (see colour scheme below). **g**, Two Ag atoms (dark green) on each face could be excised from the cluster to create  $Ag_4S_5$  capping mount structures, leaving a cubic  $Ag_{20}^{2+}$  core. The distance between the two Ag atoms at the bottom of the mount and the nearest Ag atoms of the  $Ag_{20}$  core is 2.83 Å, resulting in strong mount-to-core coupling. **h**, An alternative  $Ag_2S_5$  capping structure can be visualized as a sawhorse-shaped mount that straddles the dark-green Ag atoms of the intact dodecahedral  $Ag_{32}$  core. The  $Ag_2S_5$  mount is better defined than the  $Ag_4S_5$  mount (see **g**) because its Ag atoms are separated by a larger distance ( $> 3.1$  Å) from the nearest atoms of the  $Ag_{32}$  core, resulting in a weaker mount-to-core interaction. Colour scheme: grey, carbon; orange, oxygen; blue, exterior silver atoms in the mounts; gold, bridging sulphur atoms in the mounts. For interatomic distances, see Supplementary Information.

core with icosahedral symmetry (Fig. 3c, d) whose compact geometric shape contributes to the cluster stability; interestingly, a hollow core has been recently put forward theoretically for another thiol-protected Ag cluster<sup>25</sup>. The 20 atoms of the outer core occupy two distinct environments. Eight Ag atoms within the dodecahedral outer core define the vertices of a cube (light green in Fig. 3), the faces of which contain

the remaining 12 Ag atoms in pairs (dark green in Fig. 3) and are capped in such a way as to create an overall octahedral shape for the particle (Fig. 3f). Namely, four sulphur atoms from the *p*-MBA ligands are located on each face of the cube, such that the 24 sulphurs define a slightly distorted rhombicuboctahedron, an Archimedean solid (Fig. 3e). Each face then receives an additional Ag<sub>2</sub>S group to complete the inorganic part of the structure and the octahedral shape.

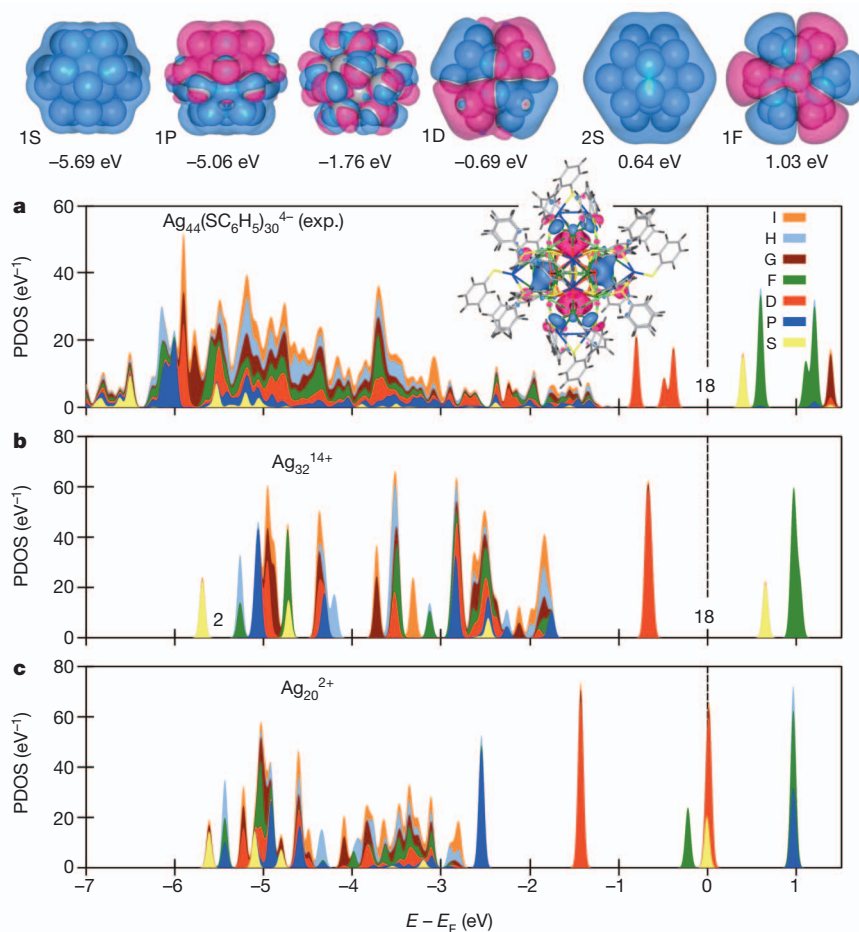
The capping units are complex three-dimensional structures unlike anything seen in gold clusters (one-dimensional gold-thiolate chains)<sup>8–12</sup> or in silver thiolate materials (two-dimensional silver-thiolate sheets)<sup>26</sup>. They can be viewed most simply as Ag<sub>2</sub>S<sub>5</sub> ‘mounts’, with four S atoms acting as legs that connect it to the Ag<sub>32</sub> core. (We borrow the term ‘mount’ from the decorative arts, where it refers to a three-dimensional decoration added to an object.) These four S atoms are bridged by a pair of Ag atoms, which are in turn bridged by a terminal S atom. Each sawhorse-shaped mount straddles a pair of Ag atoms (dark green in Fig. 3) of the intact Ag<sub>32</sub> core. Altogether, six such Ag<sub>2</sub>S<sub>5</sub> mounts comprise the entire layer protecting the compact, quasi-spherical Ag<sub>32</sub> core. The two relatively exposed Ag atoms<sup>27</sup> on each side of the six mounts can acquire effective protection from the coordinating solvent, consistent with experimental observations. If this protection is lost, the clusters can polymerize to form larger plasmonic Ag nanoparticles.

Further insight into the bonding and electronic structure of the Ag<sub>44</sub>(*p*-MBA)<sub>30</sub><sup>4–</sup> cluster was gained through extensive first-principles calculations<sup>27</sup>. Figure 4 shows the projected densities of states (PDOS, see Supplementary Information for details) calculated via density functional theory (DFT) based on the experimental configuration of the entire cluster (Fig. 3a); the PDOS reflects the angular momenta (*l*) symmetries of the cluster’s orbitals<sup>28</sup>, denoted as S, P, D,

F... (corresponding to *l* = 0, 1, 2, 3,...). The first outstanding feature observed is the relatively large energy gap between the highest occupied (HOMO) and the lowest unoccupied (LUMO) molecular orbitals,  $\Delta_{\text{HL}} = 0.78$  eV, conferring stability on the cluster and endowing it with resistance to chemical attack. The first calculated optically allowed transition occurs at 0.98 eV, however, from the 1D HOMO (here, ‘1’ is the principal quantum number) to the 1F LUMO+1 orbital, in fair agreement with the measured onset of optical absorption ( $\leq 1.1$  eV, see Fig. 1).

The wavefunctions of the cluster exhibit both localized and delocalized character. The localized states are derived from the atomic Ag 4*d* electrons and are located at the middle of the energy spectrum, whereas the delocalized states are derived from the atomic Ag 5*s* electrons and can be found near the top and bottom of the electronic spectrum<sup>28</sup>. These delocalized cluster states can be assigned angular momentum symmetries following the electronic cluster shell model<sup>28–30</sup>, with a (superatom) Aufbau rule:  $1S^2 | 1P^6 | 1D^{10} | 2S^2 | 1F^{14} | \dots$ . Here, the vertical lines denote shell-closures, which are associated with magic numbers: that is, closed-shell electronic structures are accompanied by the opening of a stabilizing energy gap. The number of electrons not engaged in bonding to sulphur (thiolates) is given by the electron count  $n^* = \nu N_{\text{Ag}} - N_{\text{L}} - Z$ , where  $\nu = 1$  is the valence for Ag (5*s*<sup>1</sup>),  $N_{\text{Ag}}$  is the number of Ag atoms,  $N_{\text{L}}$  is the number of anionic thiolate ligands, and  $Z$  is the overall cluster charge. For the case of Ag<sub>44</sub>(*p*-MBA)<sub>30</sub><sup>4–</sup>, the electron count is  $n^* = 18$ , which corresponds to the stable superatom with the Aufbau shell filling  $1S^2 | 1P^6 | 1D^{10}$ .

The 18-electron superatom shell-closure and accompanying energy gap stabilization enables a deeper insight into the structure of the protective silver-thiolate layer. Two alternative motifs for the aforementioned capping-ligand mounts can be constructed, as shown in



**Figure 4 | Projected densities of states and orbital images.**

**a**, Potential densities of states (PDOS) calculated for the Ag<sub>44</sub>(SC<sub>6</sub>H<sub>5</sub>)<sub>30</sub><sup>4–</sup> cluster with all atoms at the X-ray determined positions (see Methods and Supplementary Information). Different colours correspond to the various angular momentum contributions S, P, D, F, G, H and I, as shown on the right. The Fermi energy  $E_F$  is the energy in the middle of the HOMO–LUMO gap,  $\Delta_{\text{HL}} = 0.78$  eV. The inset shows an image of the HOMO one-dimensional superatom orbital superimposed on the structure of the cluster; different colours of the orbital (blue and pink) corresponding to different signs of the wavefunction. The 18-electron gap is marked. **b**, PDOS of the Ag<sub>32</sub><sup>14+</sup> core, as extracted from the measured structure, corresponding to the Ag<sub>2</sub>S<sub>5</sub><sup>3–</sup> mount motif. Selected superatom orbitals are shown at the top of the figure, with the energies and angular momenta marked. The 18-electron gap  $\Delta_{\text{HL}} = 1.29$  eV is noted. **c**, PDOS of the Ag<sub>20</sub><sup>2+</sup> core, as extracted from the measured structure, corresponding to the Ag<sub>4</sub>S<sub>5</sub><sup>–</sup> mount motif. Note the absence of an 18-electron gap.



Fig. 3. The cluster can be decomposed into six  $\text{Ag}_4(p\text{-MBA})_5^-$  mount units and a cubic  $\text{Ag}_{20}^{2+}$  core (Fig. 3g), where the formal charges result from the valences assigned to the silver atoms (1+) and  $p\text{-MBA}$  (thiolate) ligands (1−). Alternatively, the cluster can be decomposed into six  $\text{Ag}_2(p\text{-MBA})_5^{3-}$  mounts and a quasi-spherical  $\text{Ag}_{32}^{14+}$  core (Fig. 3h). In both cases, the cluster cores contain 18 Ag 5s electrons.

Electronic structure calculations on the two proposed cores, as extracted from the X-ray determined structure, provide a way to differentiate these two competing motifs. The PDOS of the  $\text{Ag}_{32}^{14+}$  cluster core (Fig. 4b) bears a great similarity to that of the entire cluster, including a large HOMO–LUMO gap (Fig. 4a) and a marked degeneracy of the  $1\text{D}^{10}$  superatom orbitals, reflecting an approximate spherical symmetry of the effective potential governing the motion of the delocalized electrons of the cluster core (the corresponding superatom orbital shapes are shown at the top of Fig. 4). In contrast, similar analysis for the  $\text{Ag}_{20}^{2+}$  cluster core results in a spectrum that differs considerably from that of the complete particle, and in particular does not exhibit a gap (Fig. 4c), reflecting strong coupling of this core to the  $\text{Ag}_4(p\text{-MBA})_5$  mounts. The  $\text{Ag}_2(p\text{-MBA})_5$  mounts can therefore be thought of as the operative capping unit and  $\text{Ag}_{32}$  as the natural choice for the core.

## METHODS SUMMARY

**Synthetic methods.**  $\text{M}_4\text{Ag}_{44}(p\text{-MBA})_{30}$  clusters are prepared in essentially quantitative yield by a simple three-step procedure: (1) generation of a  $\text{Ag}(\text{I})\text{-}p\text{-MBA}$  precursor; (2) reduction of the precursor to the product; and (3) removal of by-products. Afterward, the pure substance is protonated.

(1) Aqueous  $\text{AgNO}_3$  is combined with ethanolic  $p\text{-MBA}$  in excess to form the insoluble  $\text{Ag}(\text{I})\text{-}p\text{-MBA}$  precursor. The pH is then adjusted to 9 with  $\text{CsOH}$  to solubilize the precursor and is further adjusted to 12 to stabilize the final cluster product. (2) Aqueous  $\text{NaBH}_4$  is then added dropwise with stirring and is allowed to incubate for an hour. The final dark red solution yields the  $\text{M}_4\text{Ag}_{44}(p\text{-MBA})_{30}$  product with the carboxylates as alkali metal salts. (3) The clusters are separated from the reaction mixture by precipitation with DMF. The carboxylates are then protonated with acetic acid until the cluster completely dissolves in DMF, yielding the stoichiometric  $\text{M}_4\text{Ag}_{44}(p\text{-MBA})_{30}$  final product, which is isolated by precipitation.

**X-ray crystallography.** Data were collected for an  $80\text{ }\mu\text{m} \times 70\text{ }\mu\text{m} \times 50\text{ }\mu\text{m}$  crystal at 150 K with a Bruker Apex Duo diffractometer ( $\text{CuK}\alpha = 1.54178\text{ }\text{\AA}$ ) equipped with an APEX II CCD detector. The structure was solved and refined using the Bruker SHELXTL software package, with space group  $R\text{--}3c$ . All eight crystallographically independent Ag atoms were obtained by direct methods and all remaining non-hydrogen atoms were located with subsequent difference Fourier techniques. The refinement converged to  $R_1 = 5.2\%$  with a maximum resolution of  $0.83\text{ }\text{\AA}$ . The highest residual electron density was  $1.036\text{ e }\text{\AA}^{-3}$ .

**Computational methods.** The VASP-DFT package was used, with a plane-wave basis, kinetic energy cut-off of 400 eV, PAW pseudopotentials<sup>31</sup>, and the PW91 generalized gradient approximation for exchange-correlation<sup>32,33</sup>. In structural optimizations, convergence was achieved for forces  $<0.001\text{ eV }\text{\AA}^{-1}$ . Calculations were performed for the X-ray determined  $\text{Ag}_{44}(\text{SC}_6\text{H}_5)_{30}^{4-}$  structure (hydrogens were added and their positions were relaxed, average  $d(\text{C}\text{--}\text{H}) = 1.09\text{ }\text{\AA}$ ).

Received 30 March; accepted 1 August 2013.

Published online 4 September 2013.

- Heiz, U. & Landman, U. (eds) *Nanocatalysis* (Springer, 2007).
- Anker, J. N. *et al.* Biosensing with plasmonic nanosensors. *Nature Mater.* **7**, 442–453 (2008).
- Jin, R. *et al.* Controlling anisotropic nanoparticle growth through plasmon excitation. *Nature* **425**, 487–490 (2003).
- Maier, S. A. *et al.* Plasmonics—a route to nanoscale optical devices. *Adv. Mater.* **13**, 1501–1505 (2001).
- Noginov, M. A. *et al.* Demonstration of a spaser-based nanolaser. *Nature* **460**, 1110–1112 (2009).
- Atwater, H. A. & Polman, A. Plasmonics for improved photovoltaic devices. *Nature Mater.* **9**, 205–213 (2010).
- Arvizu, R. R. *et al.* Intrinsic therapeutic applications of noble metal nanoparticles: past, present and future. *Chem. Soc. Rev.* **41**, 2943–2970 (2012).
- Jadzinsky, P. D., Calero, G., Ackerson, C. J., Bushnell, D. A. & Kornberg, R. D. Structure of a thiol monolayer-protected gold nanoparticle at 1.1 Å resolution. *Science* **318**, 430–433 (2007).
- Heaven, M. W., Dass, A., White, P. S., Holt, K. M. & Murray, R. W. Crystal structure of the gold nanoparticle  $[\text{N}(\text{C}_6\text{H}_{17})_4][\text{Au}_{25}(\text{SCH}_2\text{CH}_2\text{Ph})_{18}]$ . *J. Am. Chem. Soc.* **130**, 3754–3755 (2008).

- Zhu, M., Aikens, C. M., Hollander, F. J., Schatz, G. C. & Jin, R. Correlating the crystal structure of a thiol-protected  $\text{Au}_{25}$  cluster and optical properties. *J. Am. Chem. Soc.* **130**, 5883–5885 (2008).
- Qian, H., Eckenhoff, W. T., Zhu, Y., Pintauer, T. & Jin, R. Total structure determination of thiolate-protected  $\text{Au}_{38}$  nanoparticles. *J. Am. Chem. Soc.* **132**, 8280–8281 (2010).
- Zeng, C. *et al.* Total structure of the golden nanocrystal  $\text{Au}_{36}(\text{SR})_{24}$ . *Angew. Chem. Int. Edn* **51**, 13114–13118 (2012).
- Williams, R. *The Geometrical Foundation of Natural Structure* (Dover, New York, 1979).
- Krätschmer, W., Lamb, L. D., Fostiropoulos, K. & Huffman, D. R. Solid  $\text{C}_{60}$ : a new form of carbon. *Nature* **347**, 354–358 (1990).
- Ebbesen, T. W. & Ajayan, P. M. Large-scale synthesis of carbon nanotubes. *Nature* **358**, 220–222 (1992).
- Bethune, D. S. *et al.* Cobalt-catalysed growth of carbon nano-tubes with single-atomic-layer walls. *Nature* **363**, 605–607 (1993).
- Murray, C. B., Norris, D. J. & Bawendi, M. G. Synthesis and characterization of nearly monodisperse  $\text{CdE}$  ( $\text{E} = \text{S}, \text{Se}, \text{Te}$ ) semiconductor nanocrystallites. *J. Am. Chem. Soc.* **115**, 8706–8715 (1993).
- Brust, M., Walker, M., Bethell, D., Schiffrin, D. J. & Whyman, R. Synthesis of thiol-derivatised gold nanoparticles in a two-phase liquid-liquid system. *J. Chem. Soc. Chem. Commun.* **1994**, 801–802 (1994).
- Bakr, O. M. *et al.* Silver nanoparticles with broad multiband linear optical absorption. *Angew. Chem. Int. Edn* **48**, 5921–5926 (2009).
- Harkness, K. M. *et al.*  $\text{Ag}_{44}(\text{SR})_{30}^{4-}$ : a silver–thiolate superatom complex. *Nanoscale* **4**, 4269–4274 (2012).
- Shichibu, Y. *et al.* Extremely high stability of glutathione-protected  $\text{Au}_{25}$  clusters against core etching. *Small* **3**, 835–839 (2007).
- Cathcart, N. & Kitaev, V. Silver nanoclusters: single-stage scalable synthesis of mono-disperse species and their chiro-optical properties. *J. Phys. Chem. C* **114**, 16010–16017 (2010).
- Dharmaratne, A. C., Krick, T. & Dass, A. Nanocluster size evolution studied by mass spectrometry in room temperature  $\text{Au}_{25}(\text{SR})_{18}$  synthesis. *J. Am. Chem. Soc.* **131**, 13604–13605 (2009).
- Jana, N. R., Gearheart, L. & Murphy, C. J. Seeding growth for size control of 5–40 nm diameter gold nanoparticles. *Langmuir* **17**, 6782–6786 (2001).
- Chakraborty, I. *et al.* The superstable 25-kDa monolayer protected silver nanoparticle: measurements and interpretation as an icosahedral  $\text{Ag}_{152}(\text{SCH}_2\text{CH}_2\text{Ph})_{60}$  cluster. *Nano Lett.* **12**, 5861–5866 (2012).
- Dance, I. G. The structural chemistry of metal thiolate complexes. *Polyhedron* **5**, 1037–1104 (1986).
- Herron, N., Calabrese, J. C., Farneth, W. E. & Wang, Y. Crystal structure and optical properties of  $\text{Cd}_{32}\text{S}_{14}(\text{SC}_6\text{H}_5)_{36}\cdot\text{DMF}_4$ , a cluster with a 15 Å CdS core. *Science* **259**, 1426–1428 (1993).
- Yoon, B. *et al.* Size-dependent structural evolution and chemical reactivity of gold clusters. *ChemPhysChem* **8**, 157–161 (2007).
- Knight, W. D. *et al.* Electronic shell structure and abundances of sodium clusters. *Phys. Rev. Lett.* **52**, 2141–2143 (1984).
- Yannouleas, C. & Landman, U. in *Recent Progress in Orbital-Free Density Functional Theory* (eds Wesolowski, T. A. & Wang, Y. A.) 203–250 (World Scientific, 2013).
- Kresse, G. & Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **59**, 1758–1775 (1999).
- Perdew, J. P. in *Electronic Structure of Solids '91* (eds Ziesche, P. & Eschrig, H.) 11–20 (Akademie, 1991).
- Perdew, J. P. *et al.* Atoms, molecules, solids, and surfaces: applications of the generalized gradient approximation for exchange and correlation. *Phys. Rev. B* **46**, 6671–6687 (1992); Erratum. *Phys. Rev. B* **48**, 4978–4978 (1993).

Supplementary Information is available in the online version of the paper.

**Acknowledgements** The work at the University of Toledo was supported by NSF grants CBET-0955148 and CRIF-0840474 as well as by the Wright Center for Photovoltaics Innovation and Commercialization and the School of Solar and Advanced Renewable Energy. The work of B.Y. and U.L. was supported by the Office of Basic Energy Sciences of the US Department of Energy under contract no. FG05-86ER45234 and in part by a grant from the Air Force Office of Scientific Research. Computations were made at the GATECH Center for Computational Materials Science. We acknowledge F. Stellacci for discussions and the College of Natural Sciences and Mathematics Instrumentation Center at the University of Toledo for the use of X-ray diffraction instrumentation.

**Author Contributions** T.P.B. conceived, directed and analysed all experimental research except for mass spectrometry, which W.P.G. directed and analysed, and X-ray diffraction, which K.K. directed and analysed. A.D., B.E.C. and B.M.M. performed all experimental work except for mass spectrometry, which J.G. performed. All computations and theoretical analyses were done by B.Y., R.N.B. and U.L. All authors contributed to the preparation of the manuscript.

**Author Information** The X-ray crystallographic coordinates have been deposited in the Cambridge Crystallographic Data Centre with CCDC number 949240. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to T.P.B. (Terry.Bigioni@utoledo.edu).

# Recent global-warming hiatus tied to equatorial Pacific surface cooling

Yu Kosaka<sup>1</sup> & Shang-Ping Xie<sup>1,2,3</sup>

Despite the continued increase in atmospheric greenhouse gas concentrations, the annual-mean global temperature has not risen in the twenty-first century<sup>1,2</sup>, challenging the prevailing view that anthropogenic forcing causes climate warming. Various mechanisms have been proposed for this hiatus in global warming<sup>3–6</sup>, but their relative importance has not been quantified, hampering observational estimates of climate sensitivity. Here we show that accounting for recent cooling in the eastern equatorial Pacific reconciles climate simulations and observations. We present a novel method of uncovering mechanisms for global temperature change by prescribing, in addition to radiative forcing, the observed history of sea surface temperature over the central to eastern tropical Pacific in a climate model. Although the surface temperature prescription is limited to only 8.2% of the global surface, our model reproduces the annual-mean global temperature remarkably well with correlation coefficient  $r = 0.97$  for 1970–2012 (which includes the current hiatus and a period of accelerated global warming). Moreover, our simulation captures major seasonal and regional characteristics of the hiatus, including the intensified Walker circulation, the winter cooling in northwestern North America and the prolonged drought in the southern USA. Our results show that the current hiatus is part of natural climate variability, tied specifically to a La-Niña-like decadal cooling. Although similar decadal hiatus events may occur in the future, the multi-decadal warming trend is very likely to continue with greenhouse gas increase.

Daily mean carbon dioxide at Mauna Loa of Hawaii exceeded 400 parts per million (p.p.m.) for the first time in May 2013. Whereas greenhouse gas increase has been shown to cause the centennial trend of global temperature rise since the industrial revolution<sup>7</sup>, global temperature has remained flat for the past 15 years (Extended Data Fig. 1). Two schools of thought exist regarding the cause of this hiatus in global warming: one suggests a slowdown in radiative forcing due to the stratospheric water vapour<sup>3</sup>, the rapid increase of stratospheric and tropospheric aerosols<sup>4,5</sup>, and the solar minimum around 2009 (ref. 5); and the other considers the hiatus to be part of natural variability, especially influenced by a La-Niña-like cooling in the tropical Pacific<sup>6</sup>.

A quantitative method is necessary to evaluate the relative importance of these mechanisms. Adding to the confusion surrounding the global warming hiatus, record heat waves hit Russia (summer 2010) and the USA (July 2012), and Arctic sea ice reached record lows in 2007 and 2012 (Extended Data Fig. 1). Discovering the causes of these regional climate changes requires a dynamic approach. Here we used an advanced climate model that takes radiative forcing and tropical Pacific sea surface temperature (SST) as inputs. The simulated global-mean temperature is in excellent agreement with observations, showing that the decadal cooling of the tropical Pacific causes the current hiatus. Our dynamic-model-based attribution has a distinct advantage over the empirical approach<sup>2,5</sup> in that it reveals seasonal and regional aspects of the hiatus.

Three sets of experiments were performed, using the Geophysical Fluid Dynamics Laboratory coupled model version 2.1 (ref. 8). The

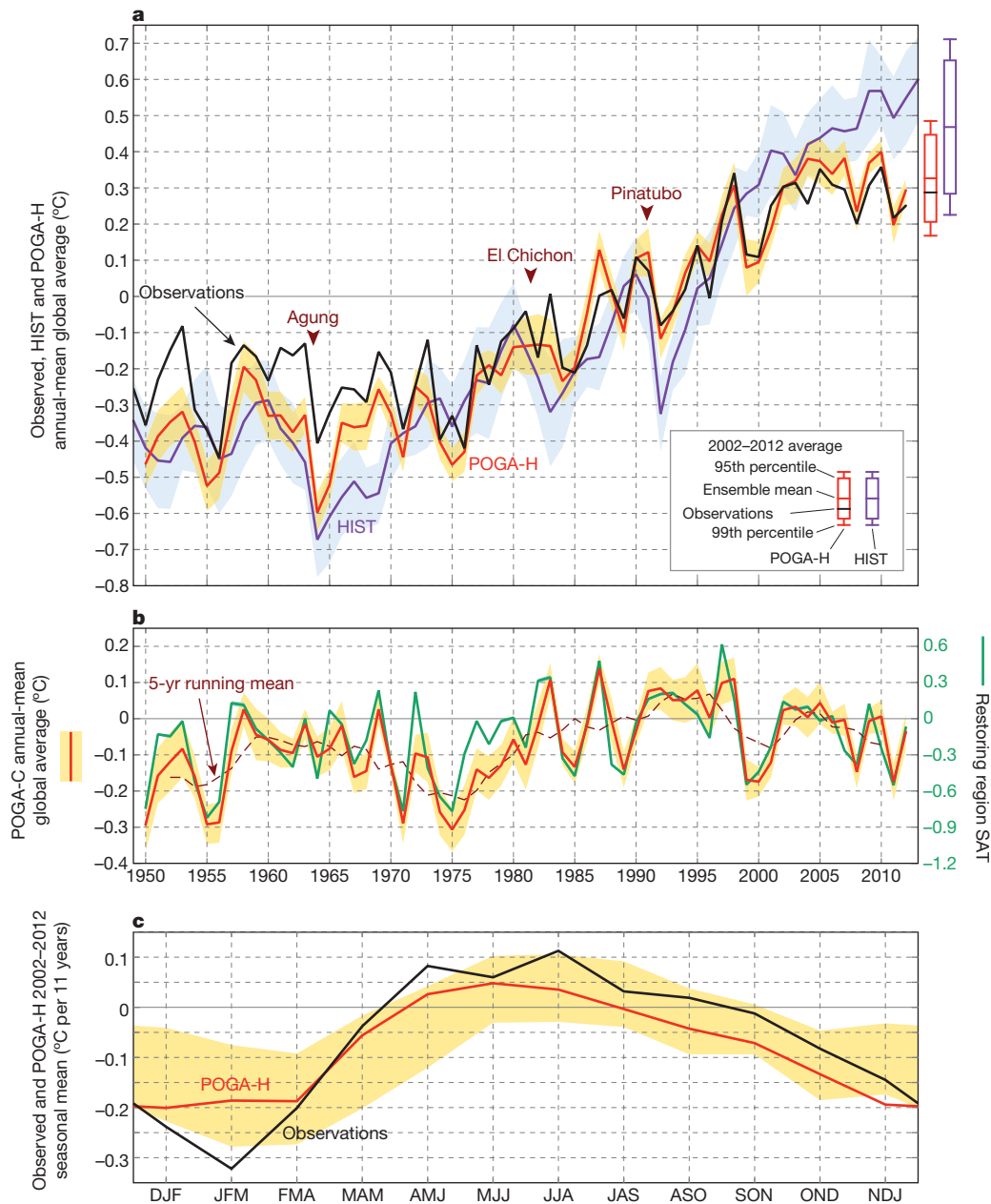
historical (HIST) experiment is forced with observed atmospheric composition changes and the solar cycle. In Pacific Ocean–Global Atmosphere (POGA) experiments, SST anomalies in the equatorial eastern Pacific (8.2% of the Earth's surface) follow the observed evolution (see Methods). In POGA-H, the radiative forcing is identical to HIST, and in the POGA control experiment (POGA-C) it is fixed at the 1990 value. Outside the equatorial eastern Pacific, the atmosphere and ocean are fully coupled and free to evolve.

Figure 1 compares the observed and simulated global near-surface temperature. In HIST, the annual-mean temperature keeps rising in response to the increased radiative forcing, with expanding departures from observations for the recent decade (Fig. 1a). POGA-H reproduces the observed record well (Extended Data Table 1). For a 43-year period after 1970 when equatorial Pacific SST data are more reliable, the correlation with observations is  $r = 0.97$  for POGA-H, due largely to the long-term trend ( $r = 0.90$  for HIST). Detrended, POGA-H still reproduces the observations at  $r = 0.70$ , whereas it falls to 0.26 in HIST. POGA-C illustrates the tropical control of the global temperature with constant radiative forcing, with the global-mean temperature closely following tropical Pacific variability (Fig. 1b). The global-mean surface air temperature (SAT) changes by 0.29 °C in response to a 1 °C SAT anomaly over the equatorial eastern Pacific. For the recent decade, the decrease in tropical Pacific SST has lowered the global temperature by about 0.15 °C compared to the 1990s (Fig. 1b), opposing the radiative forcing effect and causing the hiatus. Likewise an El-Niño-like trend in the tropics<sup>9</sup> accelerated the global warming from the 1970s to late 1990s<sup>10</sup> (Extended Data Table 1).

The POGA experimental design has been used to study the global teleconnections of the interannual El Niño/Southern Oscillation (ENSO)<sup>11,12</sup>. Here we present a novel application of POGA and demonstrate its ability to simulate the observed decadal modulations of the global warming trend, including the peculiar hiatus. Our results show that the two-parameter (radiative forcing and tropical Pacific SST) system is remarkably good at reproducing the observed global-mean temperature record, better than the HIST results with radiative forcing alone. In individual HIST realizations, hiatus events feature decadal La-Niña-like cooling in the tropical Pacific<sup>6</sup> (Extended Data Fig. 2), but POGA-H enables a direct year-by-year comparison with the observed time series of global temperature, not just the statistics from unconstrained coupled runs.

We focused on trends over the recent 11 years from 2002 to 2012, to avoid the strong 1997/98 El Niño event and the following three-year La Niña events. Net downward radiation and ocean heat content in POGA-H have continued to increase during the global SAT hiatus<sup>6</sup> (Extended Data Fig. 3). The SAT hiatus is confined to the cold season<sup>13</sup> (seasons refer to those for the Northern Hemisphere hereafter), with a decadal cooling trend for November to April, whereas the global temperature continues to rise during summer (Fig. 1c). POGA-H reproduces this seasonal cycle of the hiatus, albeit with a somewhat reduced amplitude. Although the La-Niña-like cooling trend in the tropical

<sup>1</sup>Scripps Institution of Oceanography, University of California, San Diego, 9500 Gilman Drive MC 206, La Jolla, California 92093-0206, USA. <sup>2</sup>Physical Oceanography Laboratory and Ocean–Atmosphere Interaction and Climate Laboratory, Ocean University of China, 238 Songling Road, Qingdao 266100, China. <sup>3</sup>International Pacific Research Center, SOEST, University of Hawaii at Manoa, 1680 East West Road, Honolulu, Hawaii 96822, USA.



**Figure 1 | Observed and simulated global temperature trends.** Annual-mean time series based on observations, HIST and POGA-H (a) and on POGA-C (b). Anomalies are deviations from the 1980–1999 averages, except for HIST, for which the reference is the 1980–1999 average of POGA-H. SAT anomalies over the restoring region are plotted in b, with the axis on the right. Major volcanic eruptions are indicated in a. c, Trends of seasonal global temperature for 2002–2012 in observations and POGA-H. Shading represents 95% confidence interval of ensemble means. Bars on the right of a show the ranges of ensemble spreads of the 2002–2012 averages.

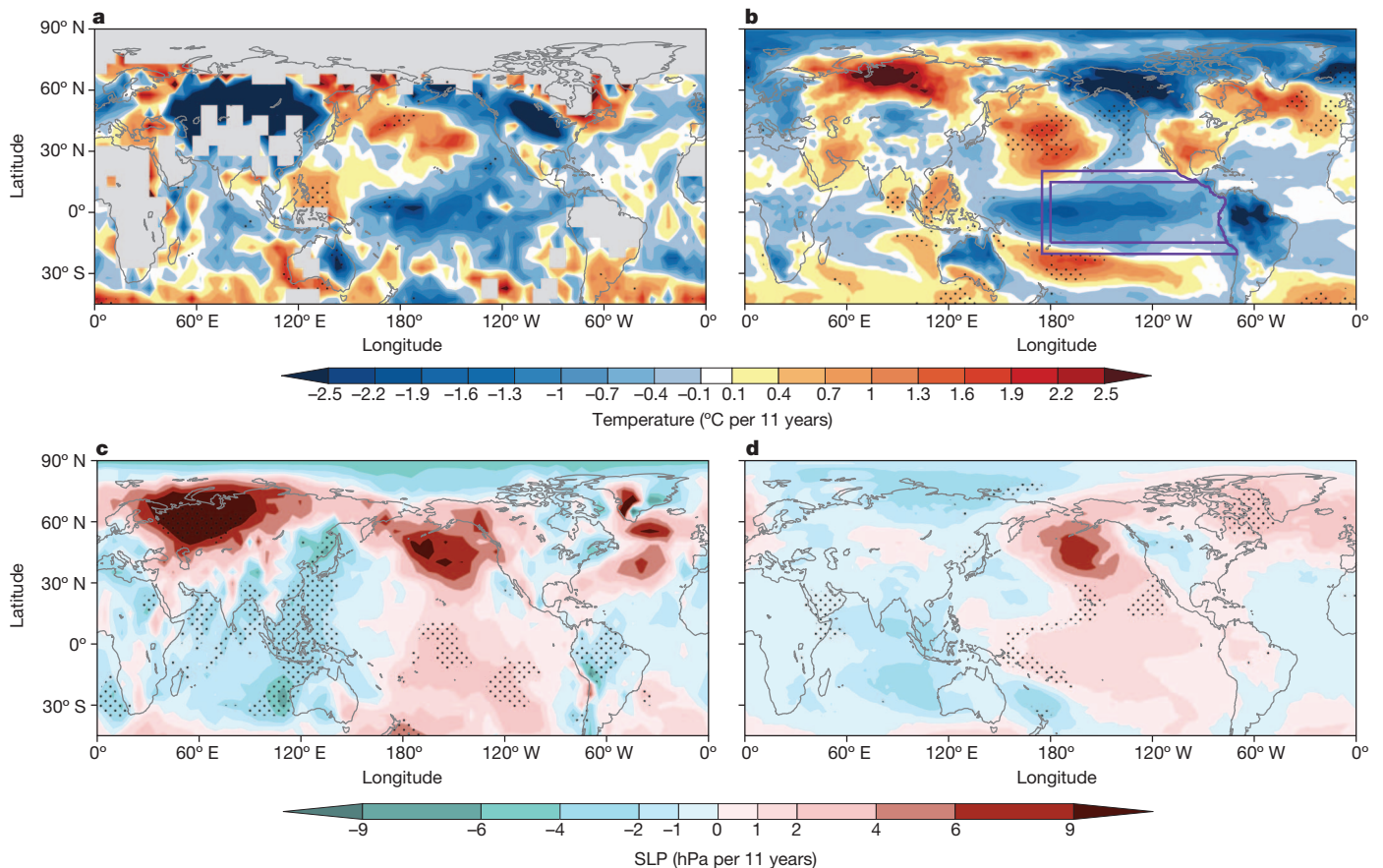
Pacific is similar in winter and in summer (Extended Data Fig. 4a), stationary/transient eddies, which are the dominant mechanism for meridional heat transport<sup>14</sup>, are stronger in winter than summer. As a result, the tropical cooling effect on the extratropics is most pronounced in winter (the seasonality of the temperature trend in the Southern Hemisphere extratropics is weak). The tropical influence on the Northern Hemisphere extratropics is weak during the summer, allowing the radiative forcing to continue the warming trend during the recent decade (Extended Data Fig. 4b).

This seasonal contrast is evident also in HIST. For 1970–2040, a period when the ensemble-mean global temperature shows a steady increase in HIST, the probability density function for the 11-year trend is similar in winter and in summer for tropical temperatures, with means both around 0.25 °C (Extended Data Fig. 4c). The probability density function is much broader for winter than for summer for Northern Hemisphere extratropical temperatures (Extended Data Fig. 4d). The chance of the 11-year temperature change falling below −0.3 °C is 8% for winter but only 0.7% for summer in the Northern Hemisphere extratropics (around 4% in the tropics for both seasons). The 11-fold

increase in the chance of an extratropical cooling in winter is partly because the tropical influence is stronger in winter than in summer.

We examined regional climate change associated with the hiatus. Although models project a slowdown of the Walker circulation in global warming<sup>15</sup>, the Pacific Walker cell intensified during the past decade (Fig. 2c). POGA-H captures this circulation change, forced by the SST cooling across the tropical Pacific (Fig. 2d). As in interannual ENSO, the tropical Pacific cooling excites global teleconnections in December, January and February (DJF; the season is denoted by the first letters of the months). SST changes in POGA-H are in broad agreement with observations over the Indian, South Atlantic and Pacific oceans outside the restoring domain (Fig. 2a, b). The model reproduces the weakening of the Aleutian low as the response of the Pacific–North American pattern to tropical Pacific cooling<sup>11</sup> (Fig. 2c, d). As a result, the SAT change over North America is well reproduced, including a pronounced cooling in the northwest of the continent. The model fails to simulate the SAT and sea-level pressure (SLP) changes over Eurasia, suggesting that they are due to internal variability unrelated to tropical forcing (Extended Data Fig. 5a and c).





**Figure 2 | Observed and simulated trend patterns in boreal winter for 2002–2012.** **a** and **b** show near-surface temperature and **c** and **d** show SLP from observations (**a** and **c**) and POGA-H (**b** and **d**) in DJF. Grey shading represents

missing values. Stippling indicates regions exceeding 95% statistical confidence. Purple boxes in **b** show the restoring region of POGA experiments.

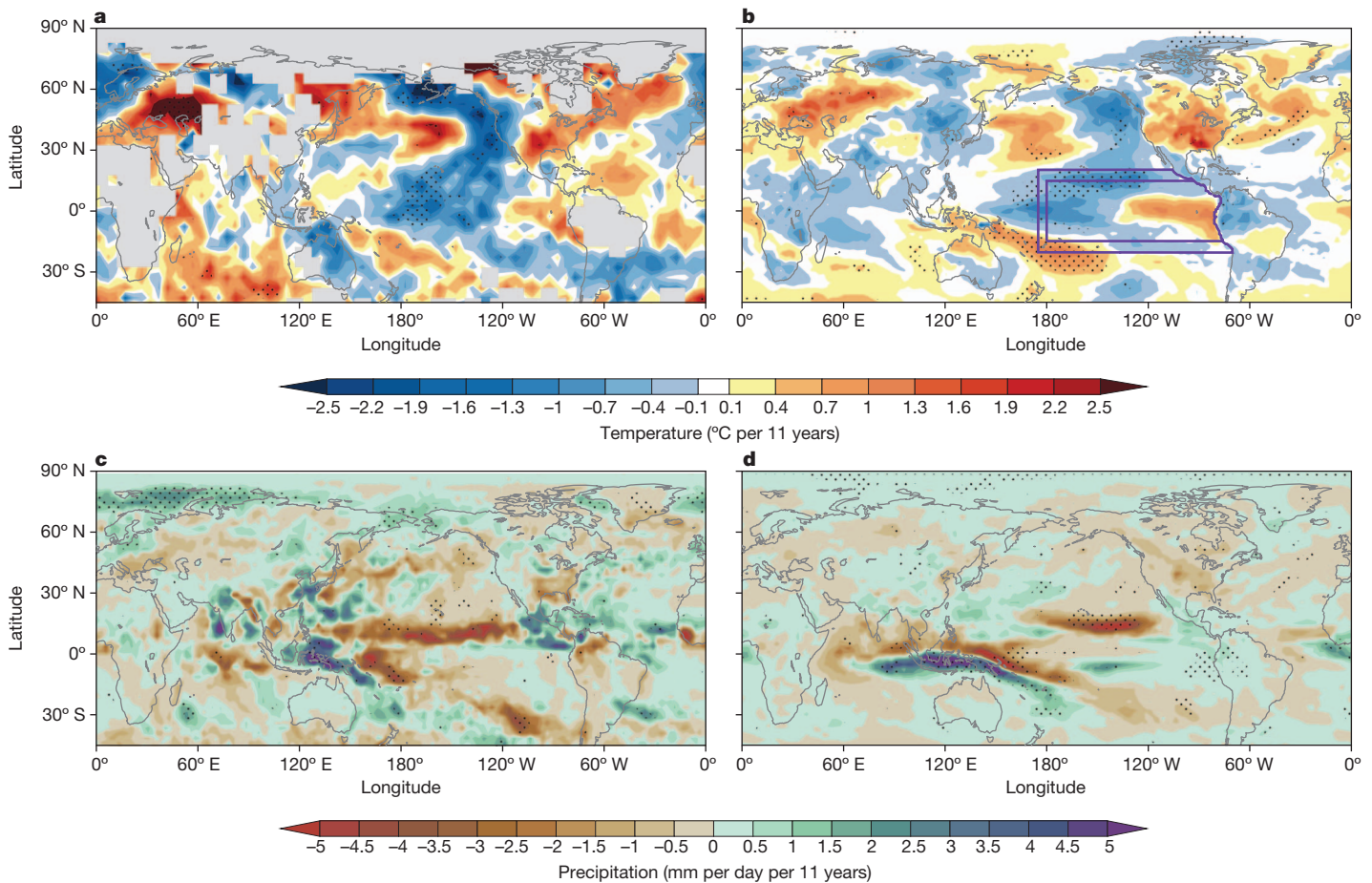
In summer, the broad agreement between simulated and observed SST remains over the Pacific (Fig. 3a, b) but the tropical influence on SAT over extratropical Eurasia and Arctic is weak (Extended Data Fig. 5b), and the increasing radiative forcing permits heatwaves to develop in Northern Hemisphere continents and Arctic sea ice to melt. (The model does not produce the record shrinkage of Arctic sea ice because of model biases and natural variability in the Arctic.) We note that the observed June, July and August (JJA) warming in western midlatitude Eurasia is much more intense than in POGA-H because heatwaves there (2003 in central Europe and 2010 in Russia) are associated with long-lasting blocking events unrelated to tropical variability<sup>16</sup>. POGA-H captures the rainfall decrease and warming over the southern USA, changes associated with prolonged droughts of record severity in Texas<sup>17</sup>. The southern USA anomalies are probably tropically forced (Extended Data Fig. 5b, d and f), for which winter–spring precipitation deficits and land-surface memory processes are probably important<sup>17</sup>. Likewise, during the epoch of accelerated global warming from the 1970s to the late 1990s, the southern USA appears as a warming hole<sup>18</sup> (local warming minimum or cooling in summer), a spatial pattern that is probably tied to tropical SST (Extended Data Fig. 6).

We present here a dynamic method for quantitative attributions of decadal modulations of global warming. By prescribing observed SST over only 8.2% of the Earth's surface, POGA-H reproduces the observed time series of global-mean temperature strikingly well, including inter-annual to decadal variability. The comparison between HIST and POGA-H indicates that the decadal cooling of the tropical Pacific is the cause of the current hiatus. In addition, POGA-H reproduces the seasonal and key regional patterns of the hiatus. The La-Niña-like cooling in the tropics affects the extratropics strongly in boreal winter, causing global cooling, a weakened Aleutian low, and an enhanced cooling over northwestern

North America among other regional anomalies. In boreal summer, in contrast, the Northern Hemisphere extratropics is largely shielded from the influence of the tropics, and the temperature continues to rise in response to the increased radiative forcing.

Whether the La-Niña-like decadal trend is internal or forced is still unclear. We note the following: (1) the tropical Pacific features pronounced low-frequency SST variability (Extended Data Fig. 7), so large that the pattern of modest forced response has not yet emerged from observations (Fig. 1b); and (2) all the climate models project a tropical Pacific warming in response to increased greenhouse gas concentrations<sup>7</sup>. We conclude that the recent cooling of the tropical Pacific and hence the current hiatus are probably due to natural internal variability rather than a forced response. If so, the hiatus is temporary, and global warming will return when the tropical Pacific swings back to a warm state. Similar hiatus events may occur in the future and are difficult to predict several years in advance owing to the limited predictability of tropical Pacific SST. We have shown that such events are accompanied by simultaneous characteristic regional patterns including an intensified Walker circulation, weakened Aleutian low and prolonged droughts in the southern USA.

Although the radiation-forced response will become increasingly important, deviations from the forced response are substantial at any given time, especially on regional scales<sup>19</sup>. We need quantitative tools—like our POGA-H—to determine the causes of regional climate anomalies<sup>17</sup>. The current hiatus illustrates the global influence of tropical Pacific SST, and a dependency of climate sensitivity on the spatial pattern of tropical ocean warming, which itself is uncertain in observations<sup>20</sup> and among models<sup>21,22</sup>. This highlights the need to develop predictive pattern dynamics constrained by observations.



**Figure 3 | Observed and simulated trend patterns in boreal summer for 2002–2012.** Same as Fig. 2, except that **a** and **b** show near-surface temperature

and **c** and **d** show precipitation from observations (**a** and **c**) and POGA-H (**b** and **d**) in JJA.

## METHODS SUMMARY

We used the Hadley Centre–Climate Research Unit combined land SAT and SST (HadCRUT) version 4.1.1.0 (ref. 23), the Hadley Centre mean SLP data set version 2 (HadSLP2, ref. 24) and monthly precipitation data from the Global Precipitation Climatology Project (GPCP) version 2.2 (ref. 25). We examined three sets of coupled model experiments based on the Geophysical Fluid Dynamics Laboratory coupled model version 2.1 (ref. 8). HIST is forced by historical radiative forcing for 1861–2005 and Representative Concentration Pathway 4.5 (RCP4.5) for 2006–2040, based on Coupled Model Intercomparison Project phase 5 (CMIP5, ref. 26). In POGA-H and POGA-C experiments, SST is restored to the model climatology plus historical anomaly by a Newtonian cooling over the deep tropical eastern Pacific. The restoring timescale is 10 days for a 50-m-deep mixed layer. Figures 2b and 3b shows the region where SST is restored; within the inner box the ocean surface heat flux is fully overridden, while in the buffer zone between the inner and outer boxes, the flux is blended with the model-diagnosed one. In POGA-H, radiative forcing is identical to HIST, whereas it is fixed at 1990 values in POGA-C. The three experiments consist of ten member runs each.

**Online Content** Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 18 June; accepted 8 August 2013.**

**Published online 28 August 2013.**

1. Easterling, D. R. & Wehner, M. F. Is the climate warming or cooling? *Geophys. Res. Lett.* **36**, L08706 (2009).
2. Foster, G. & Rahmstorf, S. Global temperature evolution 1979–2010. *Environ. Res. Lett.* **6**, 044022 (2011).
3. Solomon, S. *et al.* Contributions of stratospheric water vapor to decadal changes in the rate of global warming. *Science* **327**, 1219–1223 (2010).
4. Solomon, S. *et al.* The persistently variable “background” stratospheric aerosol layer and global climate change. *Science* **333**, 866–870 (2011).

5. Kaufmann, R. K., Kauppi, H., Mann, M. L. & Stock, J. H. Reconciling anthropogenic climate change with observed temperature 1998–2008. *Proc. Natl Acad. Sci. USA* **108**, 11790–11793 (2011).
6. Meehl, G. A., Arblaster, J. M., Fasullo, J. T., Hu, A. & Trenberth, K. E. Model-based evidence of deep-ocean heat uptake during surface-temperature hiatus periods. *Nature Clim. Change* **1**, 360–364 (2011).
7. Meehl, G. A. *et al.* in *Climate Change 2007: The Physical Science Basis* (eds Solomon, S. *et al.*) 747–845 (Cambridge University Press, 2007).
8. Delworth, T. L. *et al.* GFDL’s CM2 global coupled climate models. Part I: Formulation and simulation characteristics. *J. Clim.* **19**, 643–674 (2006).
9. Zhang, Y., Wallace, J. M. & Battisti, D. S. ENSO-like interdecadal variability: 1900–93. *J. Clim.* **10**, 1004–1020 (1997).
10. Meehl, G. A., Hu, A., Arblaster, J., Fasullo, J. & Trenberth, K. E. Externally forced and internally generated decadal climate variability associated with the Interdecadal Pacific Oscillation. *J. Clim.* <http://dx.doi.org/10.1175/JCLI-D-12-00548.1> (2013).
11. Alexander, M. A. *et al.* The atmospheric bridge: the influence of ENSO teleconnections on air–sea interaction over the global oceans. *J. Clim.* **15**, 2205–2231 (2002).
12. Lau, N.-C. & Nath, M. J. The role of the “atmospheric bridge” in linking tropical Pacific ENSO events to extratropical SST anomalies. *J. Clim.* **9**, 2036–2057 (1996).
13. Cohen, J. L., Furtado, J. C., Barlow, M., Alexeev, V. A. & Cherry, J. E. Asymmetric seasonal temperature trends. *Geophys. Res. Lett.* **39**, L04705 (2012).
14. Trenberth, K. E., Caron, J. M., Stepaniak, D. P. & Worley, S. Evolution of El Niño–Southern Oscillation and global atmospheric surface temperatures. *J. Geophys. Res.* **107**, AAC 5–1–17, <http://dx.doi.org/10.1029/2000JD000298> (2002).
15. Vecchi, G. A. *et al.* Weakening of tropical Pacific atmospheric circulation due to anthropogenic forcing. *Nature* **441**, 73–76 (2006).
16. Barriopedro, D., García-Herrera, R., Lupo, A. R. & Hernández, E. A climatology of northern hemisphere blocking. *J. Clim.* **19**, 1042–1063 (2006).
17. Hoerling, M. *et al.* Anatomy of an extreme event. *J. Clim.* **26**, 2811–2832 (2013).
18. Wang, H. *et al.* Attribution of the seasonality and regionality in climate trends over the United States during 1950–2000. *J. Clim.* **22**, 2571–2590 (2009).
19. Deser, C., Knutti, R., Solomon, S. & Phillips, A. S. Communication of the role of natural variability in future North American climate. *Nature Clim. Change* **2**, 775–779 (2012).
20. Tokinaga, H., Xie, S.-P., Deser, C., Kosaka, Y. & Okumura, Y. M. Slowdown of the Walker circulation driven by tropical Indo-Pacific warming. *Nature* **491**, 439–443 (2012).

21. DiNezio, P. N. *et al.* Climate response of the equatorial Pacific to global warming. *J. Clim.* **22**, 4873–4892 (2009).
22. Ma, J. & Xie, S.-P. Regional patterns of sea surface temperature change: a source of uncertainty in future projections of precipitation and atmospheric circulation. *J. Clim.* **26**, 2482–2501 (2013).
23. Morice, C. P., Kennedy, J. J., Rayner, N. A. & Jones, P. D. Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: the HadCRUT4 data set. *J. Geophys. Res.* **117**, D08101 (2012).
24. Allan, R. & Ansell, T. A new globally complete monthly historical gridded mean sea level pressure dataset (HadSLP2): 1850–2004. *J. Clim.* **19**, 5816–5842 (2006).
25. Adler, R. F. *et al.* The version-2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979–present). *J. Hydrometeorol.* **4**, 1147–1167 (2003).
26. Taylor, K. E., Stouffer, R. J. & Meehl, G. A. An overview of CMIP5 and the experiment design. *Bull. Am. Meteorol. Soc.* **93**, 485–498 (2012).

**Acknowledgements** We thank the Geophysical Fluid Dynamics Laboratory model developers for making the coupled model version 2.1 available and L. Xu, Y. Du, N. C. Johnson and C. Deser for discussions. The work was supported by the NSF (ATM-0854365), the National Basic Research Program of China (2012CB955600), and NOAA (NA100AR4310250).

**Author Contributions** Y.K. and S.-P.X. designed the model experiments. Y.K. performed the experiments and analysis. S.-P.X. and Y.K. wrote the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.-P.X. ([sxie@ucsd.edu](mailto:sxie@ucsd.edu)).



## METHODS

**Gridded observational data sets.** We used the Hadley Centre–Climate Research Unit combined land SAT and SST (HadCRUT) version 4.1.1.0 (<http://www.metoffice.gov.uk/hadobs/crutem4/>; ref. 23); the Hadley Centre mean SLP data set version 2 (HadSLP2, <http://www.metoffice.gov.uk/hadobs/hadslp2/>; ref. 24); and monthly precipitation data from Global Precipitation Climatology Project (GPCP) version 2.2 (<http://www.gewex.org/gpcp.html>; ref. 25). HadCRUT is compared with the SAT of POGA and HIST.

**Model experiments.** We used the Geophysical Fluid Dynamics Laboratory coupled model version 2.1 (ref. 8). The HIST, POGA-H and POGA-C experiments were formed of ten member runs each. HIST is forced by the historical radiative forcing of Coupled Model Intercomparison Project phase 5 (CMIP5, ref. 26) for 1861–2005 and extends to 2040 using Representative Concentration Pathway 4.5 (RCP4.5). The forcing includes greenhouse gases, aerosols, ozone, the solar activity cycle (repeating the cycle for 1996–2008 after 2009) and land use.

In POGA experiments, the deep tropical eastern Pacific SST was restored to the model climatology plus historical anomaly by overriding the surface sensible heat flux to ocean ( $F$ ) with:

$$F = (1 - \alpha)F_* + \alpha(cD/\tau)(T' - T_*)$$

Here a prime refers to the anomaly, asterisks represent model-diagnosed values, and  $T$  denotes SST. The reference temperature anomaly  $T'$  is based on Hadley Centre Ice and SST version 1 (HadISST1, <http://www.metoffice.gov.uk/hadobs/hadisst/>; ref. 27). The model anomaly is the deviation from the climatology of a 300-year control experiment.  $c$  is the specific heat of sea water,  $D = 50$  m is the typical depth of the ocean-mixed layer, and  $\tau = 10$  days is the restoring timescale. Figures 2b and 3b show the region where SST is restored: a weight  $\alpha = 1$  within the inner box, linearly reduced to zero in the buffer zone from the inner to the outer

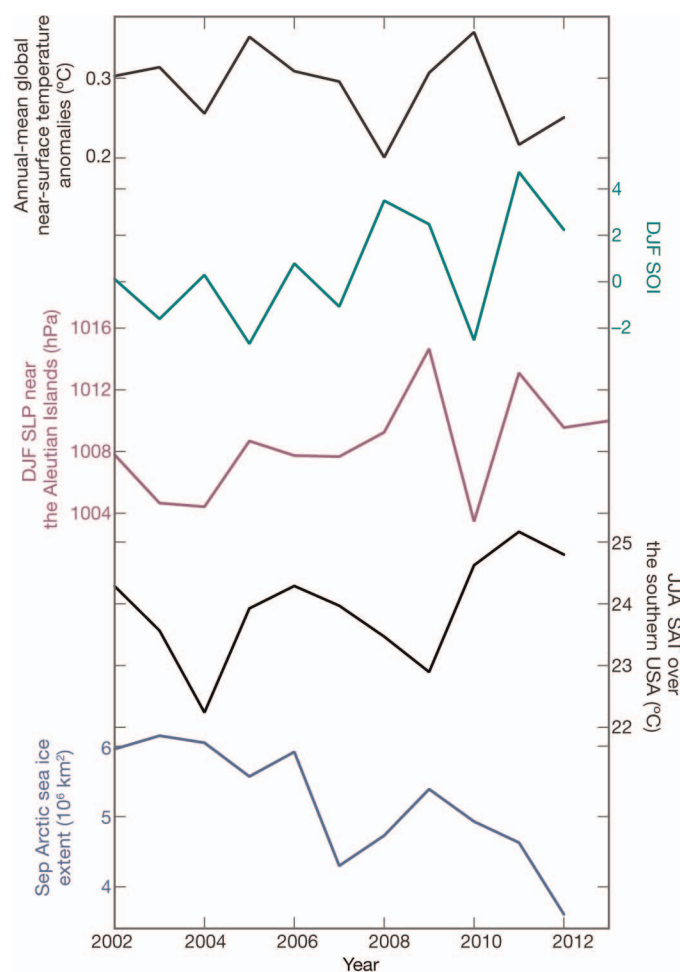
boxes. In POGA-H, radiative forcing is identical to HIST, whereas it is fixed at 1990 values in POGA-C.

**Trend estimates.** Trends were calculated as the Sen median slope<sup>28</sup>. For observed surface temperature, trends were calculated for grid boxes where data are available for more than 80% of years with at least one month per season. The Mann–Kendall test was performed for the statistical significance of trends shown in Figs 2 and 3 and Extended Data Fig. 6, and Student's  $t$ -test was applied for the significance of composited differences or anomalies of trends (Extended Data Table 1 and Extended Data Fig. 2). In Extended Data Fig. 4c and d, the trends were evaluated every four years for individual members of HIST, and probability density functions were plotted with a kernel density estimation and a Gaussian smoother.

**Decadal variability.** In Extended Data Figs 5 and 7, the Lanczos low-pass filter with a half-power frequency of eight years was applied to extract decadal variability. Extended Data Fig. 5 shows decadal anomalies obtained from a regression analysis, with their statistical significance tested with the Student's  $t$ -statistic.

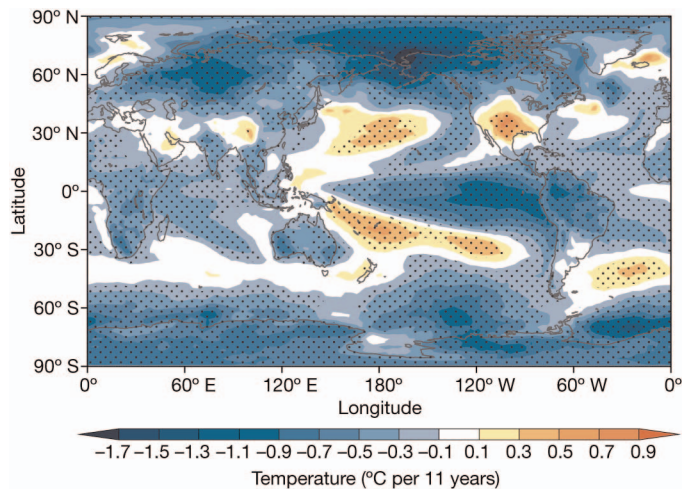
**Other observational data sets.** For Extended Data Fig. 1, we also used the Southern Oscillation Index (<http://www.cpc.ncep.noaa.gov/data/indices/>; ref. 29) and the US National Snow and Ice Data Center Arctic sea ice extent ([http://nsidc.org/data/seaice\\_index/](http://nsidc.org/data/seaice_index/); ref. 30).

27. Rayner, N. A. *et al.* Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.* **108**, 4407 (2003).
28. Sen, P. K. Estimates of the regression coefficient based on Kendall's tau. *J. Am. Stat. Assoc.* **63**, 1379–1389 (1968).
29. Trenberth, K. E. Signal versus noise in the Southern Oscillation. *Mon. Weath. Rev.* **112**, 326–332 (1984).
30. Fetterer, F. & Knowles, K. Sea ice index monitors polar ice extent. *Trans. AGU* **85**, 163 (2004).



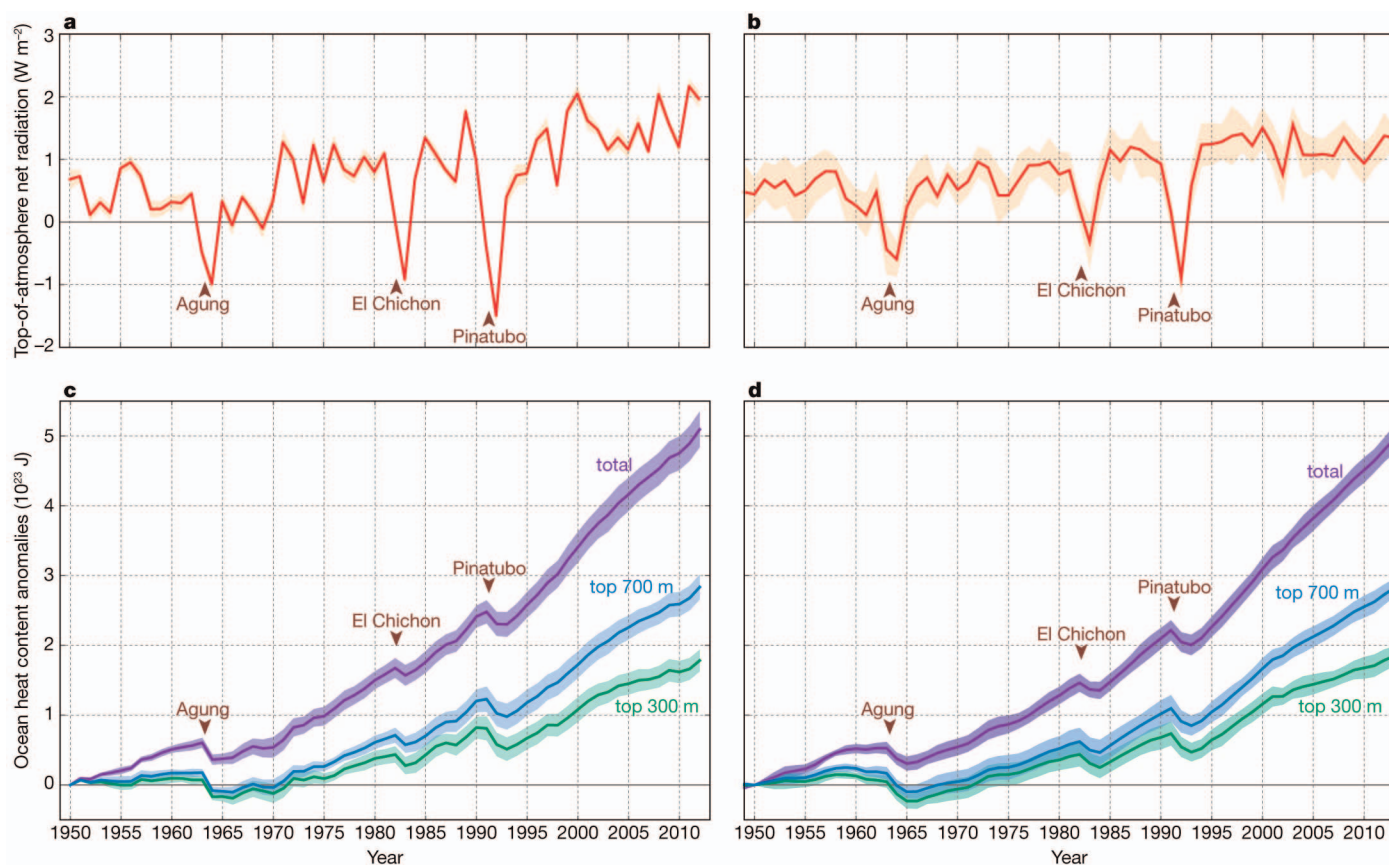
#### Extended Data Figure 1 | Observed climate indices for the recent decade.

From top to bottom are shown the annual-mean global near-surface temperature anomalies from the 1980–1999 average, the DJF Southern Oscillation Index (SOI), the DJF SLP near the Aleutian Islands (40°–60° N, 170°–120° W), the JJA SAT over the southern USA (30°–45° N, 110°–80° W), and the September (Sep) Arctic sea ice extent.



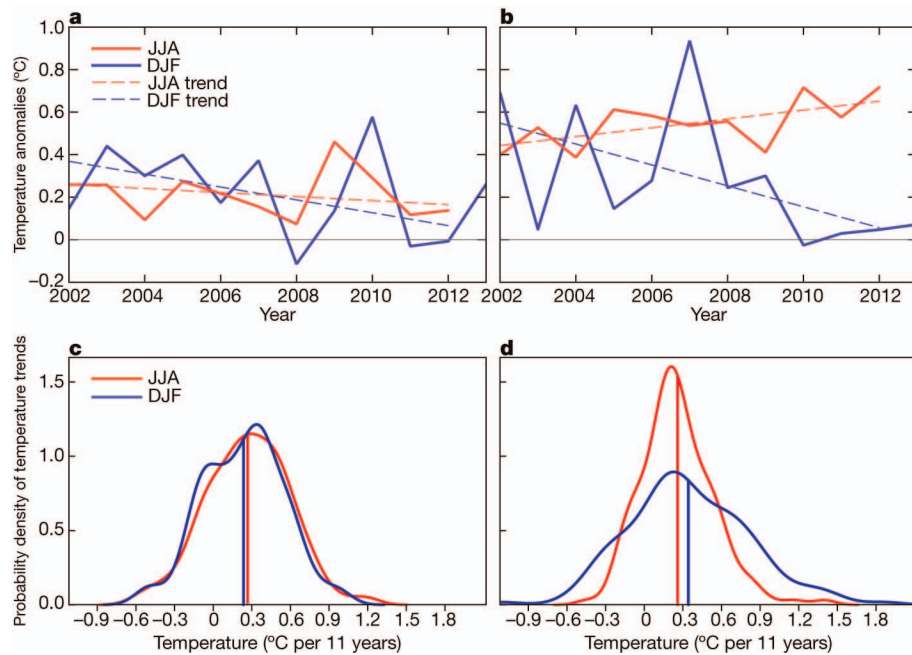
**Extended Data Figure 2 | 11-year trends of annual-mean SAT composited for 34 hiatus events in HIST.** The hiatus events are chosen for which annual-mean global SAT trends are smaller than their ensemble mean minus  $0.3^{\circ}\text{C}$  per 11 years. Stippling indicates 95% statistical confidence. Note that a typical hiatus in HIST features a La-Niña-like pattern in the tropics and SST cooling around the Aleutian Islands, patterns similar to the current hiatus event.





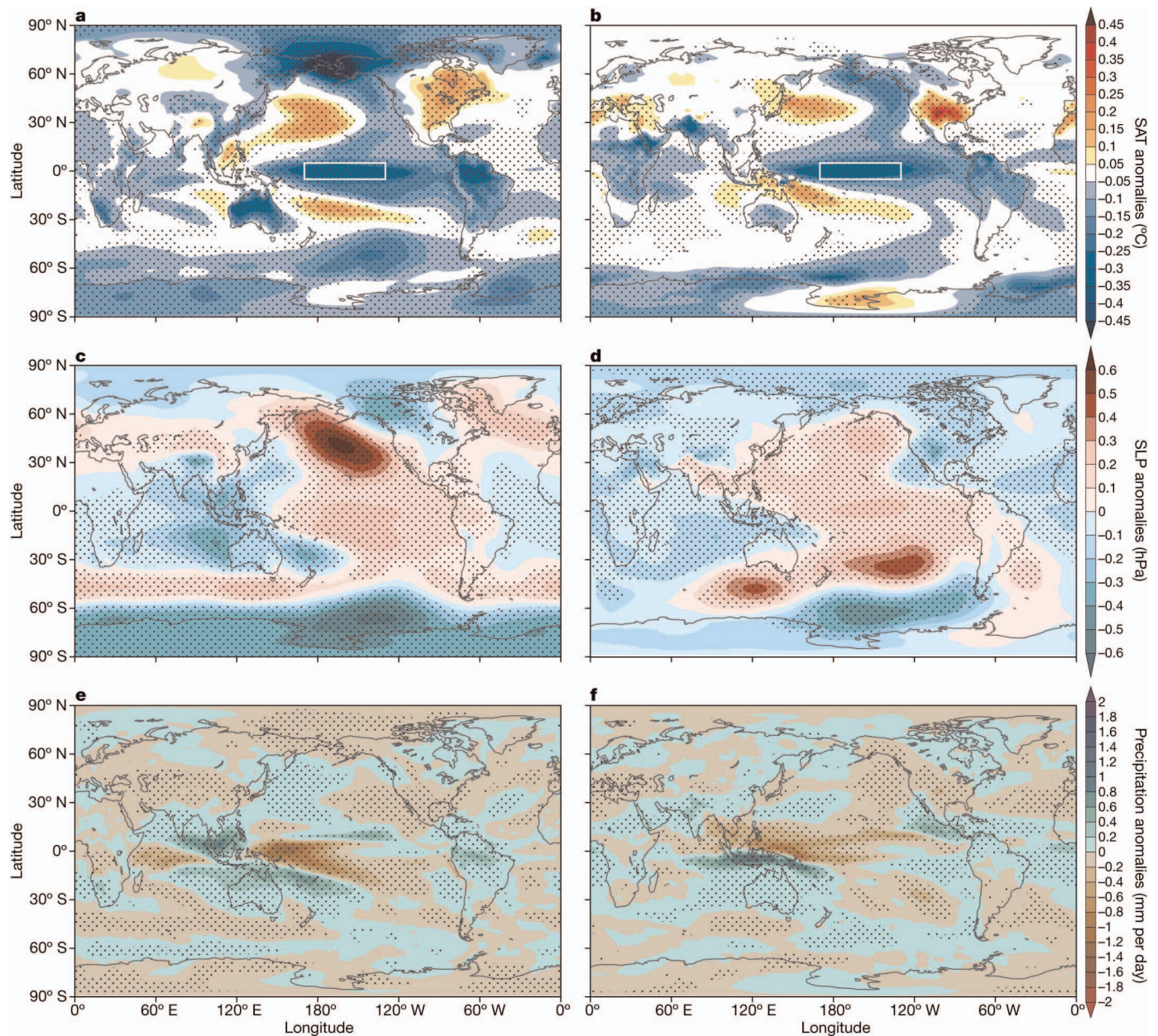
**Extended Data Figure 3 | Net radiative imbalance and ocean heat content increase in POGA-H and HIST.** a, b, Net radiative imbalance at the top of the atmosphere. Positive values indicate net energy flux into the planet. c, d, Ocean heat content deviations from 1950 values for each ensemble member. POGA-H (a and c) and HIST (b and d). Shading represents the 95%

confidence interval of ensemble means. Major volcanic eruptions are indicated. The radiative imbalance has remained positive and ocean heat content has kept increasing for the recent decade in both of the experiments. We note that the energy budget is not closed in POGA-H.



**Extended Data Figure 4 | Seasonal dependency of regional temperature trends.** **a, b,** Observed temperature anomalies (solid) and their trends (dashed) for the recent decade (°C). **c, d,** Probability density functions (curves) and means (vertical lines) of 11-year SAT trends in HIST for 1971–2040. Temperature has been averaged over the tropics (20°S–20°N; panels **a** and **c**)

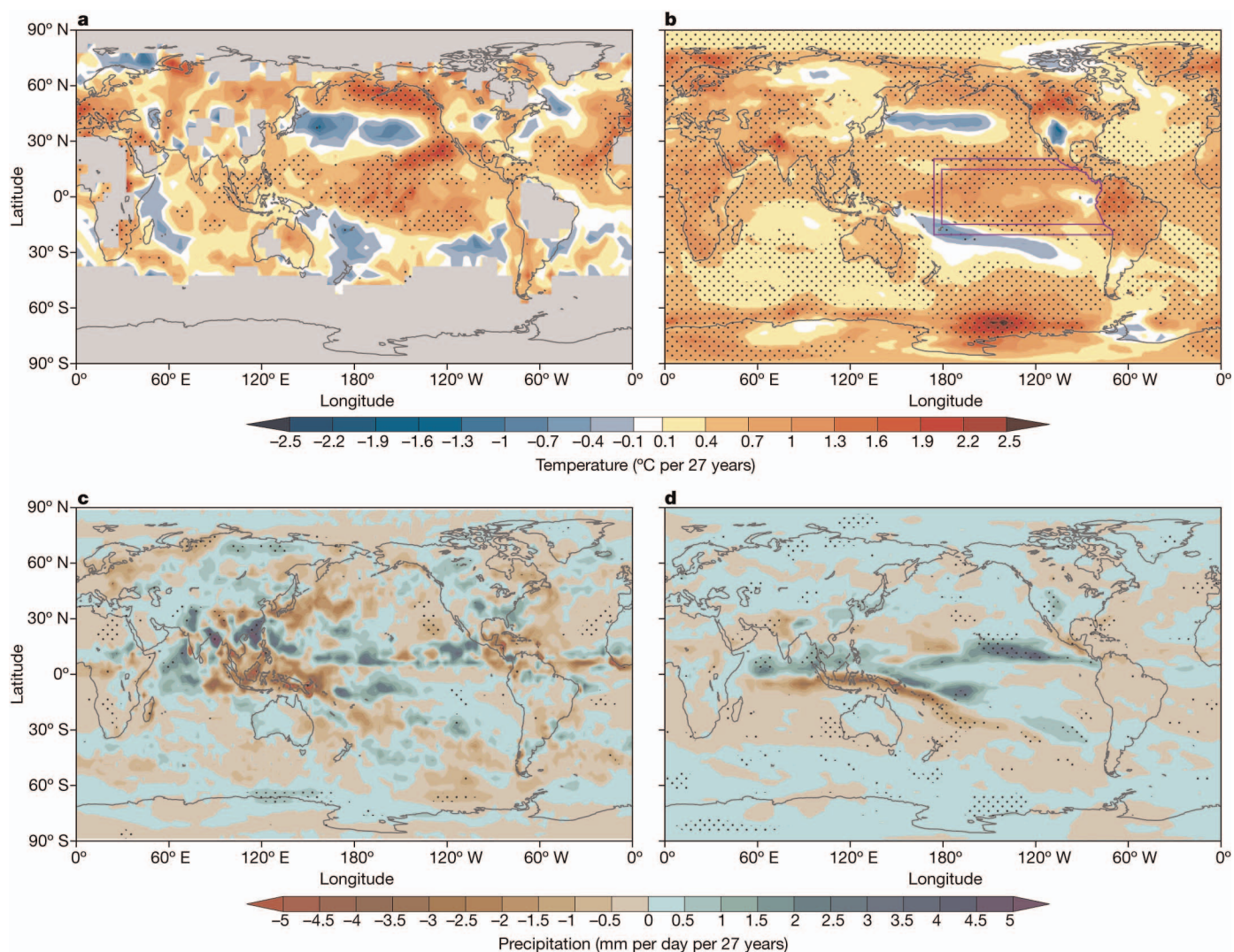
and the northern extratropics (20°–90°N; panels **b** and **d**) for JJA (red) and DJF (blue). Note that the northern extratropics features a larger probability density function spread in winter than summer, in contrast to a high similarity in the tropics. The winter spread is also greater in the extratropics than the tropics, whereas the opposite is true for summer.



**Extended Data Figure 5 | Decadal anomalies associated with SST cooling over the equatorial Pacific.** Low-pass-filtered intermember anomalies in HIST regressed against SST anomalies over the area denoted by 5°S–5°N and 170°E–130°W (white boxes in a and b). a, b, SAT; c, d, SLP; e, f, Precipitation for DJF (a, c, e) and JJA (b, d, f). The sign is flipped to show a La Niña state.

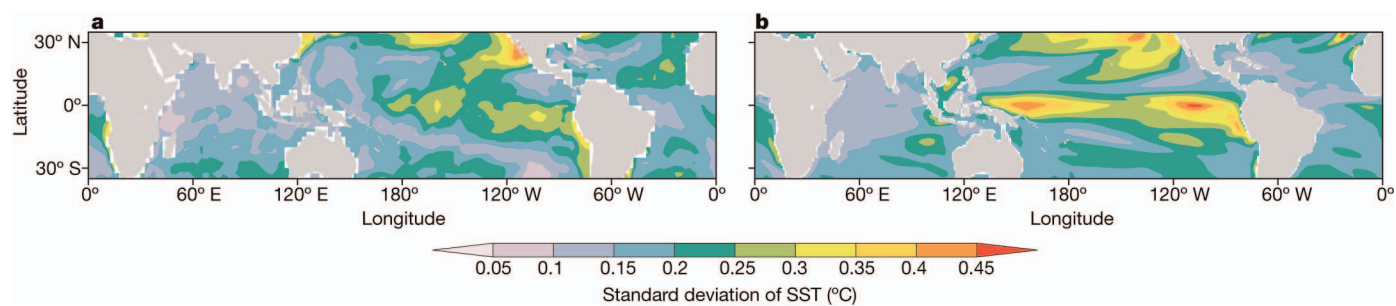
Stippling indicates 95% statistical confidence. We note that cold anomalies spread to the Arctic region in boreal winter but are restricted south of 60°N in summer. Anomalies in the tropics, the North Pacific and North America are broadly consistent with the trends for the current hiatus.





**Extended Data Figure 6 | Observed and simulated trend patterns in boreal summer for the accelerated global warming period.** Temperature (a and b) and precipitation (c and d) from observations (a and c) and POGA-H (b and d) in JJA. **a, b** and **d**, Trends for 1971–1997. **c**, The trend is evaluated for 1979–1997 and scaled to 27-year change. Stippling indicates 95% statistical

confidence. Purple boxes in **b** show the restoring region of POGA experiments. We note the widespread warming, with weak cooling in the North and South Pacific oceans and a weakened Walker circulation. POGA-H reproduces the warming hole<sup>18</sup> (warming minimum or cooling in the central USA) with slight geographical displacements due to model biases.



**Extended Data Figure 7 | Internal decadal variability in SST.** Standard deviations of annual-mean SST from observations<sup>27</sup> detrended for 1900–2012 (a) and from inter-member anomalies in HIST (b). Data were evaluated with a

decadal low-pass filter. We note that tropical variance is most pronounced in the Pacific.

**Extended Data Table 1 | Evaluation of simulations of observed global annual-mean temperature and its trend.**

Experiment	Correlation coefficient		Root mean square error (°C)		Linear trends	
	Raw	Detrended	Raw	Detrended	2002-2012 (°C per 11 years)	1971-1997 (°C per 27 years)
POGA-H	0.97	0.70	0.068	0.091	−0.01	0.55
HIST	0.90	0.26	0.149	0.143	0.19	0.41

All data are based on ensemble-mean values. Correlations and root-mean-square errors are evaluated for 1970–2012 with respect to observations. The trends are significantly different between the experiments at  $P < 0.01$  (2002–2012) and  $P < 0.05$  (1971–1997) on the basis of Student's *t*-test applied for the ensembles.



# Rapid cross-density ocean mixing at mid-depths in the Drake Passage measured by tracer release

Andrew J. Watson<sup>1†</sup>, James R. Ledwell<sup>2</sup>, Marie-José Messias<sup>1†</sup>, Brian A. King<sup>3</sup>, Neill Mackay<sup>1</sup>, Michael P. Meredith<sup>4,5</sup>, Benjamin Mills<sup>1†</sup> & Alberto C. Naveira Garabato<sup>3,6</sup>

Diapycnal mixing (across density surfaces) is an important process in the global ocean overturning circulation<sup>1–3</sup>. Mixing in the interior of most of the ocean, however, is thought to have a magnitude just one-tenth of that required to close the global circulation by the downward mixing of less dense waters<sup>4</sup>. Some of this deficit is made up by intense near-bottom mixing occurring in restricted ‘hot-spots’ associated with rough ocean-floor topography<sup>5,6</sup>, but it is not clear whether the waters at mid-depth, 1,000 to 3,000 metres, are returned to the surface by cross-density mixing or by along-density flows<sup>7</sup>. Here we show that diapycnal mixing of mid-depth (~1,500 metres) waters undergoes a sustained 20-fold increase as the Antarctic Circumpolar Current flows through the Drake Passage, between the southern tip of South America and Antarctica. Our results are based on an open-ocean tracer release of trifluoromethyl sulphur pentafluoride. We ascribe the increased mixing to turbulence generated by the deep-reaching Antarctic Circumpolar Current as it flows over rough bottom topography in the Drake Passage. Scaled to the entire circumpolar current, the mixing we observe is compatible with there being a southern component to the global overturning in which about 20 sverdrups ( $1 \text{ Sv} = 10^6 \text{ m}^3 \text{ s}^{-1}$ ) upwell in the Southern Ocean, with cross-density mixing contributing a significant fraction (20 to 30 per cent) of this total, and the remainder upwelling along constant-density surfaces. The great majority of the diapycnal flux is the result of interaction with restricted regions of rough ocean-floor topography.

Our present understanding of diapycnal mixing in the Southern Ocean is based largely on indirect inference. Observations of velocity and density ‘fine structure’, at vertical scales of order 10–100 m, have been used to infer the turbulence responsible for vertical mixing<sup>8–11</sup>, which occurs at centimetre scales and is more difficult to measure. These studies use empirical parameterizations<sup>12,13</sup> to represent the cascade of energy from larger to smaller spatial scales, induced by non-linear interactions in the internal wave field. The observations suggest that diapycnal mixing is enhanced over rough bottom topography. A mechanism thought likely to be responsible for this is the breaking of internal lee waves, generated as the deep-reaching Antarctic Circumpolar Current (ACC) encounters sea-floor features having lateral scales of a few kilometres. Accordingly, wave radiation theory has been used to estimate the energy flux into lee waves, using observation- and model-based estimates of the deep flow, and the spectral properties of such small-scale topography<sup>14,15</sup>. These studies indicate that the radiation of lee waves is sufficiently energetic to support the fine-structure estimates of diapycnal mixing. However, the fine-scale estimates are highly inferential, and the observations underpinning them are sparse. Their relevance to the integrated diapycnal mixing that affects the large-scale ocean circulation has therefore remained unclear. Furthermore, given that most of the turbulence generated by interaction with topography occurs within about 1,000 m of the bottom<sup>8,16</sup>, the question of how the water at mid-depths is returned to the surface remains unanswered. A long-standing

view is that this is accomplished by diapycnal down-mixing of buoyancy<sup>2</sup>, but the rates of mixing typically measured in the ocean’s interior are low. An alternative suggestion is that this overturning is mostly forced mechanically, by the wind acting on the surface of the Southern Ocean<sup>7,18</sup>. In the limit of no internal diapycnal mixing, modification of water densities would occur only in the surface layer, by interaction with the atmosphere.

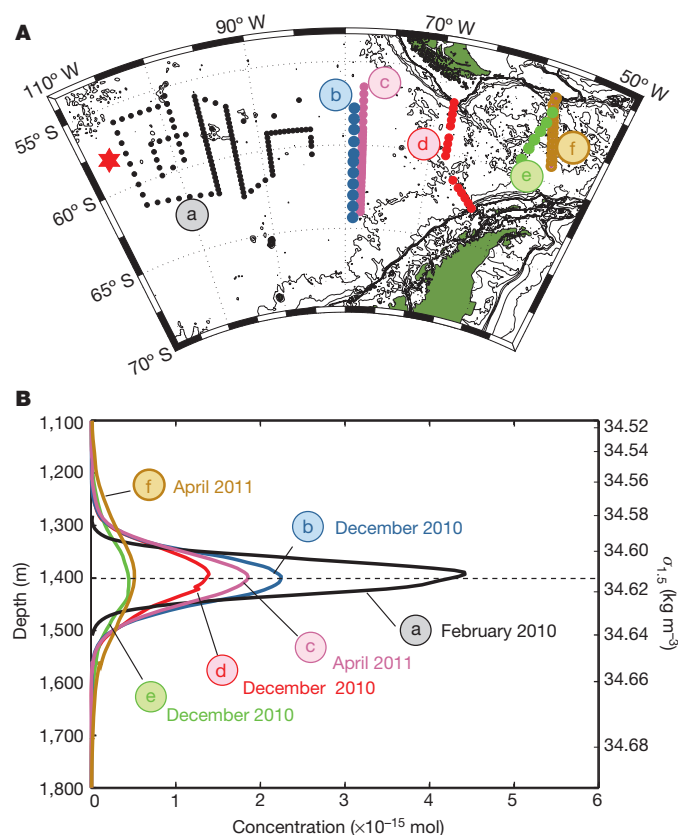
Recently, a series of microstructure measurements of the centimetre-scale turbulence responsible for vertical mixing has been conducted within, or in association with, the DIMES project (‘Diapycnal and Isopycnal Mixing Experiment in the Southern Ocean’; <http://dimes.ucsd.edu>)<sup>16,17,19</sup>. These observations are much less reliant on assumptions concerning the generation of turbulence than are the fine-structure measurements discussed above, but they are snapshots of a highly episodic and heterogeneous process, so it is difficult to scale up from these to an estimate of the mixing important on the regional or global scale. The results lend support to the idea that most mixing in the deep Southern Ocean stems from the breaking of internal lee waves, but there is a suggestion that the extent of intense turbulent mixing may have been considerably overestimated by the indirect methods. The motivation for our tracer experiment was therefore to provide an independent and unequivocal measurement of diapycnal mixing, directly relevant to larger scales and thus applicable to models of the ocean general circulation.

The tracer, 76 kg of trifluoromethyl sulphur pentafluoride ( $\text{CF}_3\text{SF}_5$ ; ref. 20), was released within 3 m depth of an isopycnal surface (with potential density, referenced to 1,500 dbar, of  $34.614 \text{ kg m}^{-3}$ ) in February 2009, as part of the DIMES project. The release location (red star in Fig. 1A) was near  $58^\circ \text{S}$ ,  $107^\circ \text{W}$ , about 2,000 km upstream of the Drake Passage, in the ACC between the Subantarctic and Polar Fronts, and at a depth of about 1,500 m (ref. 19), in the upper circumpolar deep water (UCDW) mass. The vertical and horizontal dispersion of the tracer was measured one year later, in the region between  $57^\circ$ – $62^\circ \text{S}$  and  $105^\circ$ – $85^\circ \text{W}$ , between the release site and the Drake Passage<sup>19</sup>. The vertical turbulent diffusivity integrated over that period was found to be  $(1.3 \pm 0.2) \times 10^{-5} \text{ m}^2 \text{ s}^{-1}$ , which is typical for the interior of the ocean far from boundaries. It is smaller by a factor of ten than the diffusivities of around  $10^{-4} \text{ m}^2 \text{ s}^{-1}$  that (on average) would be required to close the abyssal overturning circulation by down-mixing of buoyancy alone<sup>1,2</sup>.

In December 2010 and April 2011 two further surveys of the eastern part of the tracer patch were carried out as it flowed through the Drake Passage, yielding five meridional sections through the patch (Fig. 1). Mean vertical profiles of the sections are shown in Fig. 1B. The profiles are plotted against potential density, and also mapped into depth using a representative density–depth profile for the Drake Passage (the average of the potential density–depth profiles for the stations occupied in April 2011). The dashed horizontal line shows the ‘target’ density at which the release was made.

The surveys can be used to constrain diapycnal diffusivities both in the eastern Pacific and the Drake Passage sectors of the ACC. The

<sup>1</sup>School of Environmental Sciences, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, UK. <sup>2</sup>Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA. <sup>3</sup>National Oceanography Centre, Empress Dock, Southampton SO14 3ZH, UK. <sup>4</sup>British Antarctic Survey, High Cross, Madingley Road, Cambridge CB3 0ET, UK. <sup>5</sup>Scottish Association for Marine Science, Oban, PA37 1QA, UK. <sup>6</sup>University of Southampton, National Oceanography Centre, Empress Dock, Southampton SO14 3ZH, UK. <sup>†</sup>Present address: College of Life and Environmental Sciences, University of Exeter, Laver Building, Exeter EX4 4QE, UK.



**Figure 1 | Location of the tracer experiment and the vertical spread of the tracer during the first two years after release.** **A**, Release in February 2009 (red star) and subsequent measurements and surveys: (a), East Pacific survey, one year after release<sup>19</sup>; (b) and (c), sections near 78° W, at 1.9 years and 2.2 years after release; (d), section at the western entrance to the Drake Passage, 1.9 years after release; (e) and (f), sections at the eastern exit of the Drake Passage, 1.9 and 2.2 years after release. **B**, Mean profiles obtained from each of these locations. These are plotted in potential density space (right-hand axis,  $\sigma_{1.5}$ , referenced to 1,500 dbar). Potential density is also translated into a depth scale (left-hand axis) using a mean density–depth profile appropriate to the Drake Passage (the mean of sections (c) and (f)).

growth in the second moments of the profiles in Fig. 1 can be used to estimate diffusivity averaged over the period since release<sup>4</sup>. These are shown in Fig. 2. The estimates obtained by averaging over the path from the release point to sections (e) and (f) east of the Drake Passage are two to four times larger than those confined to the Pacific sector, indicating a large increase in the rate of cross-density mixing as the tracer is advected through the Drake Passage. This is consistent with the idea that the rough bottom topography, which dominates eastward of about 70° W, greatly influences mixing rates. The tracer has resided in the high diffusivity region for only a comparatively short time (a conservative estimate would be one quarter of the total time since release, assuming no increase in eastward velocity in the Drake Passage) so the observed broadening of profiles implies that the tracer experiences a rate of diapycnal mixing that is about an order of magnitude greater in the Drake Passage than in the eastern Pacific.

To obtain a more quantitative estimate of the mixing rate in the Drake Passage, we solved the advection–diffusion equation for the tracer on a two-dimensional (longitude and depth) domain divided into two sub-regions, east and west of 67° W, this being approximately the longitude of the Phoenix ridge, which marks the western extent of the sea-floor mountains in the Drake Passage. Vertical and horizontal diffusion and horizontal advection velocities in the two subregions were adjusted to give the best fits to the mean profiles (b) to (f) of Fig. 2 (see Methods section for full details). Uncertainties in the fitted parameters were estimated from the variation of the chi-squared statistic in parameter space

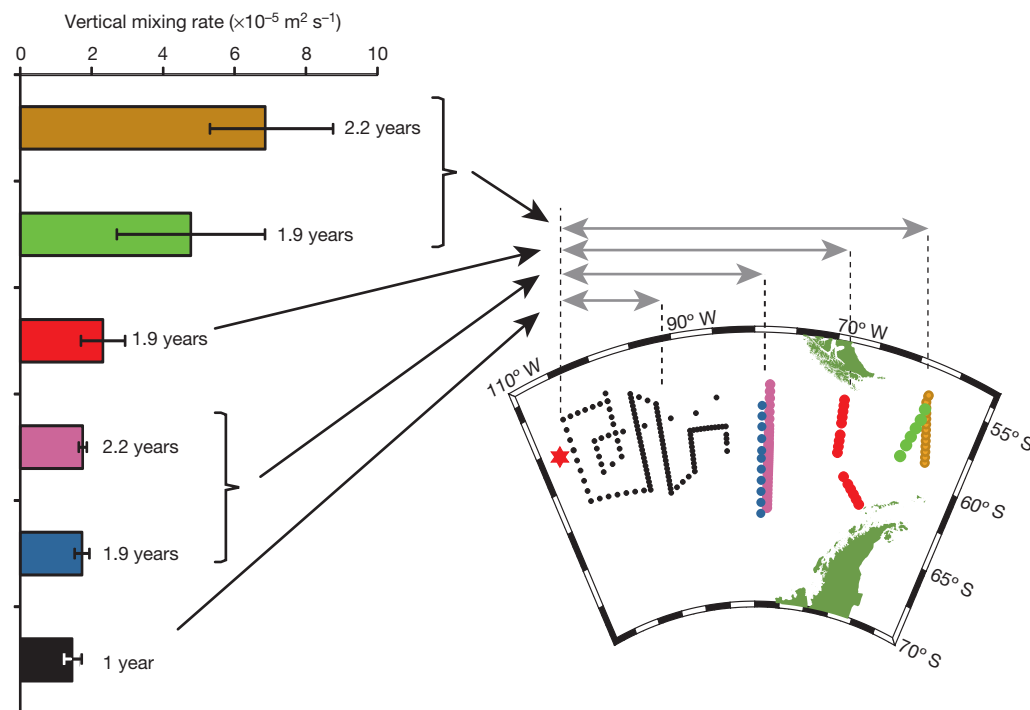
around its minimum value<sup>21</sup>. The best values estimated for the diapycnal diffusivity in the Drake Passage and in the eastern Pacific were respectively  $(3.6 \pm 0.6) \times 10^{-4} \text{ m}^2 \text{ s}^{-1}$  and  $(0.178 \pm 0.006) \times 10^{-4} \text{ m}^2 \text{ s}^{-1}$ , where uncertainties are  $2\sigma$ .

Our conclusion is that diapycnal diffusivity in the UCDW through the Drake Passage, about 2 km above the bottom, averages about 20 times the values immediately to the west in the eastern Pacific sector of the ACC. The measurements at the eastern exit of the Drake Passage were made towards the leading edge of the tracer patch. This might introduce systematic errors, biasing the mixing rate low because vertical shear has narrowed the extent of the tracer distribution, or high if more rapidly advected tracer also experiences higher vertical diffusivity than average. We estimate that such systematic errors should be contained within confidence intervals broadened by a further factor of around 1.5.

Elevated cross-density mixing has previously been observed in the vicinity of mid-ocean ridges<sup>5,6</sup>, caused by breaking of internal waves generated by bottom currents interacting with the rough topography, and the view that substantial diapycnal mixing occurs in restricted areas by such interaction is now widely accepted. Although the driver for near-bottom currents at the mid-ocean ridges is the internal tides, in the Southern Ocean the driver is mostly the deep-reaching ACC and the eddies associated with it. Globally, it has been estimated that 0.2 TW (about 20% of the wind energy put into the circulation of the ocean) may be dissipated by such interaction<sup>22</sup>, much of it in the Southern Ocean<sup>15</sup>. The great majority of the energy goes into the layer within about a kilometre of the bottom<sup>15</sup>, and the diapycnal mixing induced there is likely to be of considerable importance in the modification of the deepest water masses. This forms the return path for the lower limb of the meridional overturning circulation, supporting a flux of Antarctic bottom water of order 10 Sv (ref. 15). In contrast, the layer studied by the tracer release is 2–3 km from the bottom over most of the area, except in the restricted zones where it contacts the continental slope at the northern limit of the Drake Passage, or approaches the peaks of the highest submarine mountains. Our measurements suggest that even these limited sources of high dissipation are sufficient to produce substantially increased average mixing rates.

Profiles of turbulent dissipation made at the same time as tracer measurements during DIMES also show a substantial intensification of turbulence between the eastern Pacific and the Drake Passage above the rough topography<sup>16</sup>. The rates of mixing we measured in the UCDW seem to scale, very approximately, with the input of power into lee-wave radiation. Using the calculations of ref. 22, we find that the energy density of lee-wave generation in the ACC between the tracer release and our section (d) (see Fig. 1) averages  $0.6 \text{ mW m}^{-2}$ , whereas in the Drake Passage between sections (d) and (e)/(f) it is around  $20 \text{ mW m}^{-2}$ , and so a factor of approximately 30 greater. This is comparable to the 20-fold enhancement in diapycnal mixing rates we observe.

Using this approximate proportionality between lee-wave energies and mixing in the UCDW, we can make a rough estimate of the rates of mixing averaged around the ACC, as follows. Using the calculations of ref. 22 again, we find the average lee-wave energy dissipation under the ACC to be about  $3 \text{ mW m}^{-2}$  (the ACC being here defined as the area between the Subantarctic front in the north and the southern ACC front in the south<sup>23</sup>). Multiplying this by the relatively constant ratio between our measurements and lee-wave energies in the regions we have studied, we predict diapycnal mixing averaging  $(0.6\text{--}1) \times 10^{-4} \text{ m}^2 \text{ s}^{-1}$  for the UCDW as a whole. The mixing is concentrated over regions of rough topography such as the Drake Passage, the Scotia Sea, Crozet–Kerguelen and the southeast Indian ridge. This figure is in the range of values found by ref. 24 to be compatible with a southern component of meridional overturning of order 20 Sv, given observed temperature and salinity distributions. The mixing rate can be used in combination with data for the density structure of the Southern Ocean to estimate the contribution to this overturning due to diapycnal processes: we find



**Figure 2 | Mean diapycnal diffusivities from the point of release.** These are calculated from the second moments of the mean profiles in Fig. 1. They are averages over the times since release, and their spatial extents are indicated by the grey arrows above the inset. The approximately threefold increase in the mean when averaged over a path including the Drake Passage indicates that diffusivities increase by at least an order of magnitude east of 70° W compared to west of 70° W. We note that the water column in the Drake Passage is less

(see Supplementary Information) that this is 3–6 Sv at the density level of deepest UCDW, where the tracer was released. Our measurements support the view, therefore, that about 20%–30% of the southern component of the overturning circulation at mid-depths is sustained by diapycnal processes, with the remainder being accomplished by isopycnal transport<sup>25</sup>. Our analysis implies that almost the entire diapycnal component is generated over the restricted regions of rough bottom topography below the ACC, by the interaction of the deep-reaching current with the sea floor.

## METHODS SUMMARY

**Injection.** At the pressure and temperature of release,  $\text{CF}_3\text{SF}_5$  is a liquid immiscible with water. It was released by being forced at high pressure through fine orifices, causing it to emulsify into fine droplets that dissolved without settling appreciably in the water column. The injection package was towed slowly through the water at about 1,500 m depth, and included a conductivity–temperature–depth (CTD) instrument to allow monitoring of the density of the water, with the towing winch constantly adjusted to keep the package at the target density.

**Sampling and analysis.** Water samples were collected using conventional 10-litre Niskin bottles. Two-litre subsamples were drawn from these into ground-glass-stoppered bottles, filled from the bottom using Tygon tubing and with the sample overflowed by at least two litres to expel all water that had contacted the atmosphere. Samples were kept cold and analysed within 24 h of collection. For analysis, subsamples of about one litre were stripped of volatiles using a vacuum purge-and-trap, gas chromatography and electron capture detection. Calibration for  $\text{CF}_3\text{SF}_5$  used a gravimetrically prepared gas standard.

**Diapycnal diffusivities.** These were examined by plotting individual tracer profiles and section averages (see Fig. 1) as a function of potential density, and transforming these into depth space using a depth–density profile representative of the Drake Passage. Gaussian curves were fitted to these profiles, and average mixing rates were calculated from the width of these fitted curves. More detailed estimates used the numerical solution of the advection–diffusion equation in two dimensions, optimizing the fit to the observed average profiles ((b) to (f) in Fig. 1) as functions of diapycnal diffusivity and along-stream velocity in two subregions: east of 67° W (the Drake Passage) and west of this longitude (east Pacific).

stratified than in the eastern Pacific, and so the tracer distribution on survey (a) in Fig. 1 occupies a wider depth span when mapped in a Drake Passage depth–density profile than when used in the calculations of ref. 19. Correspondingly, the diffusivity shown here for the Pacific sector after one year is about 25% larger than that quoted in ref. 19. Error bars show 95% confidence limits, calculated from the statistics of individual profiles (see Supplementary Information for details).

**Online Content** Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 14 March; accepted 3 July 2013.

1. Munk, W. & Wunsch, C. Abyssal recipes II: energetics of tidal and wind mixing. *Deep-Sea Res.* **45**, 1977–2010 (1998).
2. Munk, W. H. Abyssal recipes. *Deep-Sea Res.* **13**, 707–730 (1966).
3. Lumpkin, R. & Speer, K. Global ocean meridional overturning. *J. Phys. Oceanogr.* **37**, 2550–2562 (2007).
4. Ledwell, J. R., Watson, A. J. & Law, C. S. Evidence for slow mixing across the pycnocline from an open-ocean tracer-release experiment. *Nature* **364**, 701–703 (1993).
5. Ledwell, J. R. *et al.* Evidence for enhanced mixing over rough topography in the abyssal ocean. *Nature* **403**, 179–182 (2000).
6. Polzin, K. L., Toole, J. M., Ledwell, J. R. & Schmitt, R. W. Spatial variability of turbulent mixing in the abyssal ocean. *Science* **276**, 93–96 (1997).
7. Toggweiler, J. R. & Samuels, B. On the ocean's large scale circulation near the limit of no vertical mixing. *J. Phys. Oceanogr.* **28**, 1832–1852 (1998).
8. Garabato, A. C. N. *et al.* Widespread intense turbulent mixing in the Southern Ocean. *Science* **303**, 210–213 (2004).
9. Kunze, E. *et al.* Global abyssal mixing inferred from lowered ADCP shear and CTD strain profiles. *J. Phys. Oceanogr.* **36**, 1553–1576 (2006).
10. Sloyan, B. M. Spatial variability of mixing in the Southern Ocean. *Geophys. Res. Lett.* **32**, L18603 (2005).
11. Wu, L., Jing, Z., Riser, S. & Visbeck, M. Seasonal and spatial variations of Southern Ocean diapycnal mixing from Argo profiling floats. *Nature Geosci.* **4**, 363–366 (2011).
12. Gregg, M. C. Scaling turbulent dissipation in the thermocline. *J. Geophys. Res.* **94**, 9686–9698 (1989).
13. Polzin, K. L., Toole, J. M. & Schmitt, R. W. Finescale parameterizations of turbulent dissipation. *J. Phys. Oceanogr.* **25**, 306–328 (1995).
14. Nikurashin, M. & Ferrari, R. Radiation and dissipation of internal waves generated by geostrophic motions impinging on small-scale topography: application to the Southern Ocean. *J. Phys. Oceanogr.* **40**, 2025–2042 (2010).
15. Nikurashin, M. & Ferrari, R. Radiation and dissipation of internal waves generated by geostrophic motions impinging on small-scale topography: theory. *J. Phys. Oceanogr.* **40**, 1055–1074 (2010).
16. St Laurent, L. *et al.* Turbulence and diapycnal mixing in Drake Passage. *J. Phys. Oceanogr.* **42**, 2143–2152 (2012).
17. Waterman, S., Naveira Garabato, A. C. & Polzin, K. L. Internal waves and turbulence in the Antarctic Circumpolar Current. *J. Phys. Oceanogr.* **43**, 259–282 (2013).



18. Marshall, J. & Speer, K. Closure of the meridional overturning circulation through Southern Ocean upwelling. *Nature Geosci.* **5**, 171–180 (2012).
19. Ledwell, J. R., St Laurent, L. C., Garton, J. B. & Toole, J. M. Diapycnal mixing in the Antarctic Circumpolar Current. *J. Phys. Oceanogr.* **41**, 241–246 (2011).
20. Ho, D. T., Ledwell, J. R. & Smethie, W. M. Use of SF<sub>5</sub>CF<sub>3</sub> for ocean tracer release experiments. *Geophys. Res. Lett.* **35**, L04602 (2008).
21. Bevington, P. R. & Robinson, D. K. *Data Reduction and Error Analysis in the Physical Sciences* (McGraw-Hill, 1992).
22. Nikurashin, M. & Ferrari, R. Global energy conversion rate from geostrophic flows into internal lee waves in the deep ocean. *Geophys. Res. Lett.* **38**, L08610 (2011).
23. Orsi, A. H., Whitworth, T. & Nowlin, W. D. On the meridional extent and fronts of the Antarctic Circumpolar Current. *Deep-Sea Res.* **42**, 641–673 (1995).
24. Zika, J. D., Sloyan, B. M. & McDougall, T. J. Diagnosing the Southern Ocean overturning from tracer fields. *J. Phys. Oceanogr.* **39**, 2926–2940 (2009).
25. Webb, D. J. & Sugimotohara, N. Oceanography—vertical mixing in the ocean. *Nature* **409**, 37 (2001).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank the officers and staff of the RV *Thomas Thompson*, RRS *James Cook* and RRS *James Clark Ross* for their assistance in making the observations at sea. We thank the UK Natural Environment Research Council and the US National Science Foundation for funding the DIMES experiment. A.J.W. thanks the Royal Society for support.

**Author Contributions** A.J.W., J.R.L., M.-J.M., M.P.M. and A.C.N.G. planned and directed the tracer experiment. M.J.M. led the chemical analysis at sea, and, together with A.J.W., J.R.L., N.M. and B.M., obtained the tracer data. A.J.W., J.R.L., M.P.M., N.M., B.A.K. and A.C.N.G. analysed the physical oceanographic data. N.M. carried out the two-dimensional model computations. M.P.M., A.J.W. and A.C.N.G. planned and directed the research cruises. A.J.W. wrote the initial draft of the paper and all authors contributed to its revision.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.J.W. ([andrew.watson@exeter.ac.uk](mailto:andrew.watson@exeter.ac.uk)).

## METHODS

**Tracer injection.** The 75 kg of trifluoromethyl sulphur pentafluoride ( $\text{CF}_3\text{SF}_5$ ) was deployed with a towed system similar to that described by ref. 26 for an injection of  $\text{SF}_6$  in the North Atlantic. The underwater components of the injection system included a CTD unit that measured temperature, conductivity and pressure, from which density was calculated; a set of water samplers for calibration; control units; two tracer reservoirs; a primer reservoir; lead-acid batteries; and two high-pressure pumps feeding eight 25- $\mu\text{m}$ -diameter ceramic orifices. In operation, the tracer was forced by the pumps through the orifices at pressures of approximately 10 MPa. These components were mounted on an aluminium frame that was made neutrally buoyant with a set of glass floats mounted on the top of the frame. The frame was attached to the termination of the CTD cable with a 2.5-m tether to remove much of the ship's motion. The CTD cable was held down with a hydroweight suspended below the termination. The system was kept within a metre (root mean squared (r.m.s.) distance) of the target density surface with a feedback system whereby the departure from the target density measured at the CTD was used to adjust the winch. A primer fluid, 2,3-dihydroperfluoropentane, was used to test the system on deck (because the tracer is a gas at normal temperature and pressure), and was also used to keep the orifices clear as the injection system was lowered to the target surface.

The tracer was injected in a crude 'x' pattern centred near  $58.1^\circ\text{S}$ ,  $106.7^\circ\text{W}$  at an average rate of 0.9 g per second. The target density was originally defined as a potential density, referenced to 1,500 dbar, of  $34.6077\text{ kg m}^{-3}$ . This density corresponded closely with the  $27.90\text{ kg m}^{-3}$  neutral density surface in the region of the injection. The average tow speed during injection was  $0.50\text{ m s}^{-1}$ , and the overall injection track length was 40 km. The initial distribution of the tracer was sampled within two weeks of release with both a towed array of samplers similar to that used in previous experiments<sup>26</sup> and with a conventional CTD/rosette system with 22 four-litre Niskin bottles. The mean vertical profile of the tracer obtained with the two methods agreed, but the measurements were more accurate for the CTD/rosette samples. Samples were transferred into syringes and analysed by the head space method<sup>27</sup>. The mean vertical distribution exhibited a single peak, approximately Gaussian, with the centre of mass 2 m below the target density surface and with r.m.s. spread about this centre of mass of 5.5 m. The slight sinking of the tracer was probably due to a density excess of the tracer plume immediately after injection, and the r.m.s. spread was probably due to turbulence behind the towed injection system.

**Sample collection and analysis technique.** The released tracer and two transient tracers, trifluorochloromethane (CFC-13) and dichlorodifluoromethane (CFC-12), were measured on board ship by a purge-and-trap gas chromatographic method with electron capture detection. The instrumentation was developed at the University of East Anglia following earlier designs<sup>20,28,29</sup>. Water sampling used an instrument package including a CTD and rosette of 24 ten-litre General Oceanics sampling bottles with metal springs, the nitrile 'O' rings of which had been decontaminated following standard chlorofluorocarbon sampling procedures. Water samples were collected from these bottles into two-litre ground-glass-stoppered bottles filled from the bottom using Tygon tubing, and overflowed with at least two litres to expel all water exposed to the air. The glass bottles were stored immersed in sea water at a temperature below  $5^\circ\text{C}$ , and analysed within 24 h. The measurement system stripped volatiles from a 1.1-litre subsample of the water. Calibration for  $\text{CF}_3\text{SF}_5$  used a gravimetrically prepared gas standard, which agreed within 10% with the calibration scale of E. Busenberg (personal communication; see also ref. 30). Precision for gas standards was  $<1\%$ , and for duplicate water samples the standard deviation averaged  $1\%$  or  $5 \times 10^{-18}\text{ mol per litre}$ , whichever was greater.

**Average profiles, second moments and uncertainty estimates.** The depth of the isopycnal on which the tracer was released changes substantially across the study area, from approximately 500 m in the south to approximately 2,000 m in the north of the Drake Passage (see Supplementary Fig. 1). To examine diapycnal transport, individual tracer profiles from the surveys were plotted against potential density, and section averages such as those in Fig. 1 were calculated by averaging along isopycnals rather than at constant depths. However, rates of diapycnal diffusivity are normally expressed in units of length squared divided by time, by analogy with molecular diffusivity, so all isopycnal averages were transformed into depth using a single representative depth–density profile, following the techniques described in ref. 26. The 'standard' density profile used for this is shown in Supplementary Fig. 2, and is the average potential density of sections (c) and (f) (see Fig. 1), intermediate between the eastern and the western Drake Passage.

Estimates of the overall diapycnal diffusivity, averaged over the path since the release point, were calculated by fitting the solution to the one-dimensional diffusion equation with constant diffusivity to the section-average concentration profiles, assuming a Gaussian initial distribution. This solution is  $c = c_m \exp(-(z - z_m)^2 / 2\sigma^2)$  where  $z$  is depth and  $c_m$ ,  $z_m$  and  $\sigma$  are fitted variables. The diffusivity  $k$  is then equal to

$(\sigma^2 - \sigma_0^2) / 2t$ , where  $t$  is the time since release and  $\sigma_0$  is the second moment of the initial distribution at time  $t = 0$ .

**Model to estimate diapycnal diffusivities.** More detailed estimates of transport parameters from the tracer distribution require a numerical model treatment. Full three-dimensional model realizations are currently in progress, but for this work, approximate estimates were made using a relatively simple two-dimensional approach. We solved a two-dimensional realization of the advection–diffusion equation for a conserved tracer  $c$

$$\frac{\partial c}{\partial t} = \nabla \cdot (k \nabla c) - u \cdot \nabla c \quad (1)$$

where  $k$  is a tensor of eddy diffusion and  $u$  is a velocity. The vertical coordinate was taken normal to isopycnal surfaces and extended 300 m around the release isopycnal, and the horizontal dimension was zonal distance along the path of the ACC and in a domain that extended from  $120^\circ\text{W}$  to  $30^\circ\text{W}$ . Eddy diffusion  $k$  was modelled with an along-stream isopycnal diffusion  $k_h$  and a diapycnal, vertical diffusion  $k_z$ , whereas velocity had an along-stream component only. Zonal dispersion was initially set by comparison to the dispersion of clusters of model floats released at 1,500 m in the CCSM model<sup>31</sup> (M. E. Maltrud and J. L. McClean, personal communication). These suggest zonal dispersion in the ACC corresponding to about  $3,000\text{ m}^2\text{ s}^{-1}$  in the first one to two years following release. Horizontal velocities were initially set by averaging zonal velocities at the target isopycnal derived from the SATGEM product for Southern Ocean dynamics<sup>32</sup>. Velocity and the two components of diffusivity were held constant within each of two sectors: west of  $67^\circ\text{W}$  (the east Pacific) and east of that point (the Drake Passage). For each sector the velocities were assumed to decrease with depth according to an exponential profile with a scale height of 1,500 m (ref. 32). Six parameters were then fitted, these being the magnitudes of the along-stream velocity,  $k_h$  and  $k_z$ , in each of the two sectors.

To optimize parameters, vertical widths of the distributions predicted by the model were compared to widths of individual profiles measured on sections (b) to (f) (see Fig. 1). Repeated runs of the model were used to map out and minimize the following cost function as a function of the six variable parameters:

$$\chi^2 = \sum_{i=1}^n \frac{(w_i - W_{b-f})^2}{\sigma_{b-f}^2} \quad (2)$$

Here,  $w_i$  are the measured widths of the individual profiles and  $W_{b-f}$  are the widths of the appropriate model profiles with which to compare them (for example, the output of the model at the time after release and longitude corresponding to the measured sections (b) to (f)).  $\sigma_{b-f}^2$  are the variances of profile widths as measured on each of those sections. Within wide limits, we found that the procedure of fitting the vertical widths of the profiles ensures that values for diapycnal diffusivity are insensitive to the choice of along-stream dispersion, and values quoted in the text use a constant value for  $k_h$  set at  $3,000\text{ m}^2\text{ s}^{-1}$ , a relatively high value that includes the effects of meridional shear dispersion<sup>33</sup>. The final fit therefore had four free parameters, rather than six.

If each profile width is considered to be a measurement of a normally distributed variate, and the model gave unbiased estimates of that variate, the expectation value  $\langle \chi^2 \rangle$  of the cost function would be the number of degrees of freedom of the estimation problem, namely  $N - q$ , where  $N$  is the number of profile widths to be fitted and  $q$  is the number of parameters that are varied to optimize the fit. For our case  $N = 75$  and  $q$  is 4, so  $\langle \chi^2 \rangle$  is 71. The increase in the cost function as the fitted parameters are varied about their optimum values can be used to estimate the uncertainty in the estimations<sup>21</sup>.

Contours of the cost function  $\chi$  as a function of diapycnal diffusivity in the Drake Passage and the east Pacific are shown in Supplementary Fig. 3. An estimate of the  $2\sigma$  uncertainty in the diffusivity can be made from the change in underlying parameters required to increase the value of  $\chi$  by two (ref. 21), and these are used for our uncertainty estimates in the main text.

Supplementary Fig. 4 shows a histogram of the widths of individual profiles on section (c) in Fig. 1, the section for which we have the largest number of profiles. The comparison shown there with a normal curve having the same mean and standard deviation suggests that the statistics are unlikely to be normal, with the upper tail in particular being elevated in comparison to a normal distribution. (However, using the Shapiro–Wilk test<sup>34</sup>, the hypothesis of normality is not rejected at  $P = 0.05$  for this data set.) The actual minimum value of the cost function was 81, higher than the expectation value for normal statistics, which also indicates some departure from normality for the data as a whole, and so the uncertainties we quote must be considered guidelines only.

26. Ledwell, J. R., Watson, A. J. & Law, C. S. Mixing of a tracer in the pycnocline. *J. Geophys. Res.* **C 103**, 21499–21529 (1998).

27. Wanninkhof, R., Ledwell, J. R. & Watson, A. J. Analysis of sulfur hexafluoride in seawater. *J. Geophys. Res.* **C 96**, 8733–8740 (1991).

28. Smethie, W. M., Schlosser, P., Bonisch, G. & Hopkins, T. S. Renewal and circulation of intermediate waters in the Canadian Basin observed on the SCICEX 96 cruise. *J. Geophys. Res. C* **105**, 1105–1121 (2000).
29. Law, C. S., Watson, A. J. & Liddicoat, M. I. Automated vacuum analysis of sulfur-hexafluoride in seawater—derivation of the atmospheric trend (1970–1993) and potential as a transient tracer. *Mar. Chem.* **48**, 57–69 (1994).
30. Busenberg, E. & Plummer, L. N. Dating groundwater with trifluoromethyl sulfurpentafluoride ( $\text{SF}_5\text{CF}_3$ ), sulfur hexafluoride ( $\text{SF}_6$ ),  $\text{CF}_3\text{Cl}$  (CFC-13), and  $\text{CF}_2\text{Cl}_2$  (CFC-12). *Wat. Resour. Res.* **44**, W02431 (2008).
31. McClean, J. L. *et al.* A prototype two-decade fully-coupled fine-resolution CCSM simulation. *Ocean Model.* **39**, 10–30 (2011).
32. Meijers, A. J. S., Bindoff, N. L. & Rintoul, S. R. Estimating the four-dimensional structure of the Southern Ocean using satellite altimetry. *J. Atmos. Ocean. Technol.* **28**, 548–568 (2011).
33. Falco, P. & Zambianchi, E. Near surface structure of the Antarctic Circumpolar Current derived from World Ocean Circulation Experiment drifter data. *J. Geophys. Res. C* **116**, C05003 (2011).
34. Shapiro, S. S. & Wilk, M. B. Analysis of variance test for normality (complete samples). *Biometrika* **52**, 591–611 (1965).

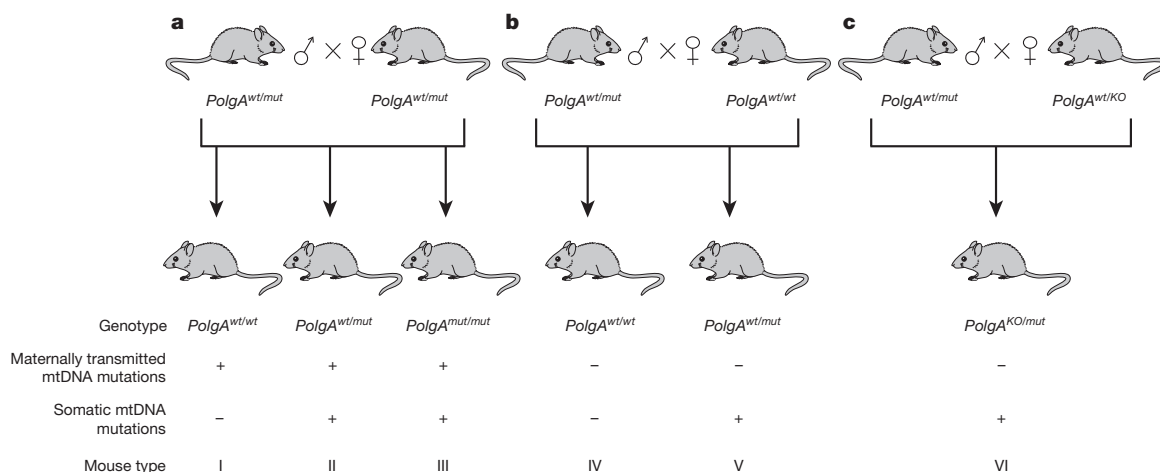


# Germline mitochondrial DNA mutations aggravate ageing and can impair brain development

Jaime M. Ross<sup>1\*</sup>, James B. Stewart<sup>2\*</sup>, Erik Hagström<sup>3</sup>, Stefan Brené<sup>4</sup>, Arnaud Mourier<sup>2</sup>, Giuseppe Coppotelli<sup>1</sup>, Christoph Freyer<sup>2,3</sup>, Marie Lagouge<sup>2</sup>, Barry J. Hoffer<sup>5</sup>, Lars Olson<sup>1</sup> & Nils-Göran Larsson<sup>2,3</sup>

Ageing is due to an accumulation of various types of damage<sup>1,2</sup>, and mitochondrial dysfunction has long been considered to be important in this process<sup>3–8</sup>. There is substantial sequence variation in mammalian mitochondrial DNA (mtDNA)<sup>9</sup>, and the high mutation rate is counteracted by different mechanisms that decrease maternal transmission of mutated mtDNA<sup>10–13</sup>. Despite these protective mechanisms<sup>14</sup>, it is becoming increasingly clear that low-level mtDNA heteroplasmy is quite common and often inherited in humans<sup>15,16</sup>. We designed a series of mouse mutants to investigate the extent to which inherited mtDNA mutations can contribute to ageing. Here we report that maternally transmitted mtDNA mutations can induce mild ageing phenotypes in mice with a wild-type nuclear genome. Furthermore, maternally transmitted mtDNA mutations lead to anticipation of reduced fertility in mice that are heterozygous for the mtDNA mutator allele (*PolgA<sup>wt/mt</sup>*) and aggravate premature ageing phenotypes in mtDNA mutator mice (*PolgA<sup>mut/mt</sup>*). Unexpectedly, a combination of maternally transmitted and somatic mtDNA mutations also leads to stochastic brain malformations. Our findings show that a pre-existing mutation load will not only allow somatic mutagenesis to create a critically high total mtDNA mutation load sooner but will also increase clonal expansion of mtDNA mutations<sup>17</sup> to enhance the normally occurring mosaic respiratory chain deficiency in ageing tissues<sup>18,19</sup>. Our findings suggest that maternally transmitted mtDNA mutations may have a similar role in aggravating aspects of normal human ageing.

We generated a series of inbred mutant mice (Fig. 1a–c) to study the role of mtDNA mutations in ageing, taking into account that *PolgA<sup>wt/mt</sup>* mice transmit low levels of mtDNA mutations through the germ line<sup>20</sup>. A standard intercross of *PolgA<sup>wt/mt</sup>* mice (Fig. 1a) was used to generate: type I mice (mice with a wild-type nuclear genome (*wt<sup>N</sup>*) containing maternally transmitted mtDNA mutations); type II mice (*PolgA<sup>wt/mt</sup>* mice containing maternally transmitted mtDNA mutations); and type III mice (*PolgA<sup>mut/mt</sup>* mice<sup>21,22</sup> containing maternally transmitted mtDNA mutations). In addition, we crossed *PolgA<sup>wt/mt</sup>* males to wild-type females (Fig. 1b) to generate: type IV mice (*wt<sup>N</sup>* mice lacking maternally transmitted mtDNA mutations) and type V mice (*PolgA<sup>wt/mt</sup>* mice lacking maternally transmitted mtDNA mutations). Finally, we crossed *PolgA<sup>wt/mt</sup>* males with females that have a heterozygous knock-out for the catalytic subunit of mitochondrial DNA polymerase<sup>23</sup> (*PolgA<sup>wt/KO</sup>*) (Fig. 1c) to generate type VI mice (a variant of mtDNA mutator mice (genotype *PolgA<sup>KO/mt</sup>*) lacking maternally transmitted mtDNA mutations). Next, we quantitatively assessed mtDNA mutation loads and found that type I–III mice contained an excess of recurring mtDNA mutations ( $P < 0.0001$ , Fisher's exact test) (Supplementary Fig. 1a, b and Supplementary Table 1) undergoing clonal expansion (Supplementary Table 2). The PCR, cloning and sequencing protocol has, on theoretical grounds, been criticized for artificially inducing mtDNA mutations<sup>24,25</sup>. We therefore used an independent method,  $\lambda$ -phage cloning of full-length mtDNA<sup>20</sup>, to analyse type I, II, III and V mice and found that the mtDNA mutation levels were very similar



**Figure 1 | Breeding to generate mice with different combinations of maternally inherited and somatic mtDNA mutations.** **a**, Intercrossing of mice heterozygous for the mtDNA mutator allele (*PolgA<sup>wt/mt</sup>*) generates type I–III mice, which inherit mtDNA mutations from their *PolgA<sup>wt/mt</sup>* mother. **b**, Crossing of *PolgA<sup>wt/mt</sup>* males and wild-type females will generate type IV–V

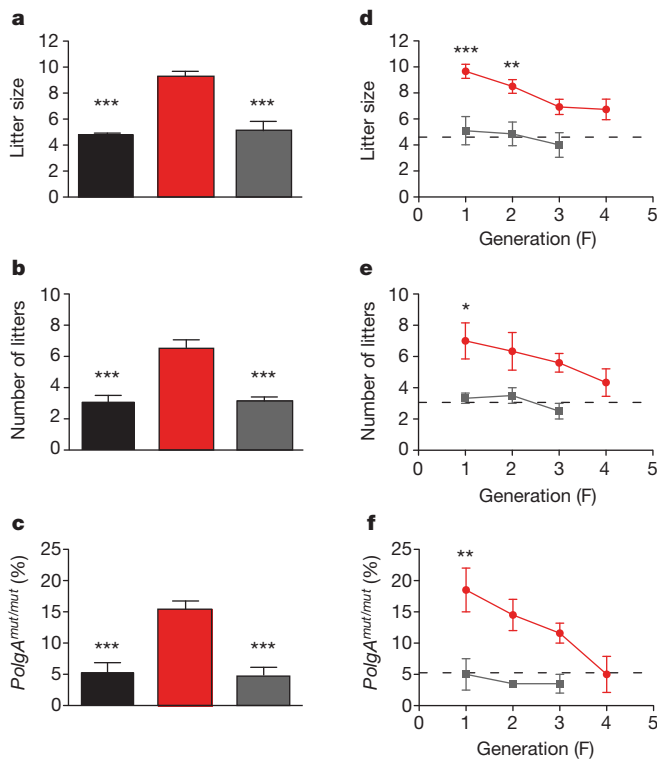
mice that lack maternally transmitted mtDNA mutations. **c**, Crossing of *PolgA<sup>wt/mt</sup>* males and heterozygous *PolgA* knockout females (*PolgA<sup>wt/KO</sup>*) will generate type VI mice that are an mtDNA mutator mouse variant lacking maternally inherited mtDNA mutations.

<sup>1</sup>Department of Neuroscience, Karolinska Institutet, Retzius väg 8, 171 77 Stockholm, Sweden. <sup>2</sup>Max Planck Institute for Biology of Ageing, Joseph-Stelzmann-Strasse 9b, 50931 Cologne, Germany. <sup>3</sup>Department of Laboratory Medicine, Karolinska Institutet, Retzius väg 8, 171 77 Stockholm, Sweden. <sup>4</sup>Department of Neurobiology, Care Sciences, and Society, KERIC, Karolinska Institutet, 171 76 Stockholm, Sweden. <sup>5</sup>Department of Neurosurgery, University Hospitals, Case Western Reserve Medical Center, 11100 Euclid Avenue, Cleveland, Ohio 44106, USA.

\*These authors contributed equally to this work.

with both methods (Supplementary Fig. 1a, c and Supplementary Table 1), demonstrating that confounding artificial mutations are not induced by the PCR, cloning and sequencing protocol.

Continuous intercrosses of  $PolgA^{wt/mut}$  animals to generate  $PolgA^{mut/mut}$  mice (Fig. 1a) produce low litter sizes (Fig. 2a, black bar), a low total number of litters per female (Fig. 2b, black bar) and less than Mendelian proportions of mtDNA mutator mice (Fig. 2c, black bar). We re-introduced wild-type mtDNA into  $PolgA^{wt/mut}$  females and a subsequent intercross gave larger litter sizes (Fig. 2a, red bar), a higher number of litters per female (Fig. 2b, red bar) and a high proportion of born  $PolgA^{mut/mut}$  mice (Fig. 2c, red bar). The reciprocal experiment, wherein wild-type males were crossed with  $PolgA^{wt/mut}$  females to obtain  $PolgA^{wt/mut}$  animals for subsequent intercrossing showed no improvement of fertility phenotypes (Fig. 2a–c, grey bars). Next, we decided to continue intercrossing  $PolgA^{wt/mut}$  animals for several generations without additional re-introductions of wild-type mtDNA and found a continuous decline of litter sizes (Fig. 2d, red line), number of litters per



**Figure 2 | Anticipation of reduced fecundity obtained by intercrossing of heterozygous mtDNA mutator mice.** **a**, Litter sizes from: standard intercrosses of  $PolgA^{wt/mut}$  mice maintained by intercrosses for several generations (black bar,  $n = 51$ ), intercrosses of  $PolgA^{wt/mut}$  mice after re-introduction of wild-type mtDNA into the females (red bar,  $n = 42$ ), intercrosses of  $PolgA^{wt/mut}$  mice obtained by crossing wild-type males to heterozygous  $PolgA^{wt/mut}$  females (grey bar,  $n = 15$ ). **b**, Number of litters per breeding pair. Crosses are as indicated in the legend to panel **a**. **c**, Proportion of live-born  $PolgA^{mut/mut}$  mice. Crosses are as indicated in the legend to panel **a**. **d**, Litter sizes after continuous intercrosses of  $PolgA^{wt/mut}$  mice. The dashed line indicates the mean litter size of a standard intercross ( $n = 51$ ). The red line indicates an intercross in which the mother in generation  $F_0$  contains re-introduced wild-type mtDNA. The mean litter sizes of  $PolgA^{wt/mut}$  mothers belonging to generations  $F_1$ – $F_4$  of this intercross are shown ( $n = 21, 21, 15$  and  $9$ , respectively). The grey line indicates an intercross in which the  $PolgA^{wt/mut}$  mice were obtained by crossing wild-type males to heterozygous mtDNA mutator females. The mean litter size of mothers belonging to generations  $F_1$ – $F_3$  of this intercross is shown ( $n = 9, 6$  and  $9$ , respectively). **e**, Number of litters after continuous intercrosses of  $PolgA^{wt/mut}$  mice. Crosses are as indicated in the legend to panel **d**. **f**, Proportion of live-born  $PolgA^{mut/mut}$  mice after continuous intercrosses of  $PolgA^{wt/mut}$  mice. Crosses are as indicated in the legend to panel **d**. One-way ANOVA (**a**–**c**) and two-way ANOVA (**d**–**f**). \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , Error bars denote  $\pm$  s.e.m.

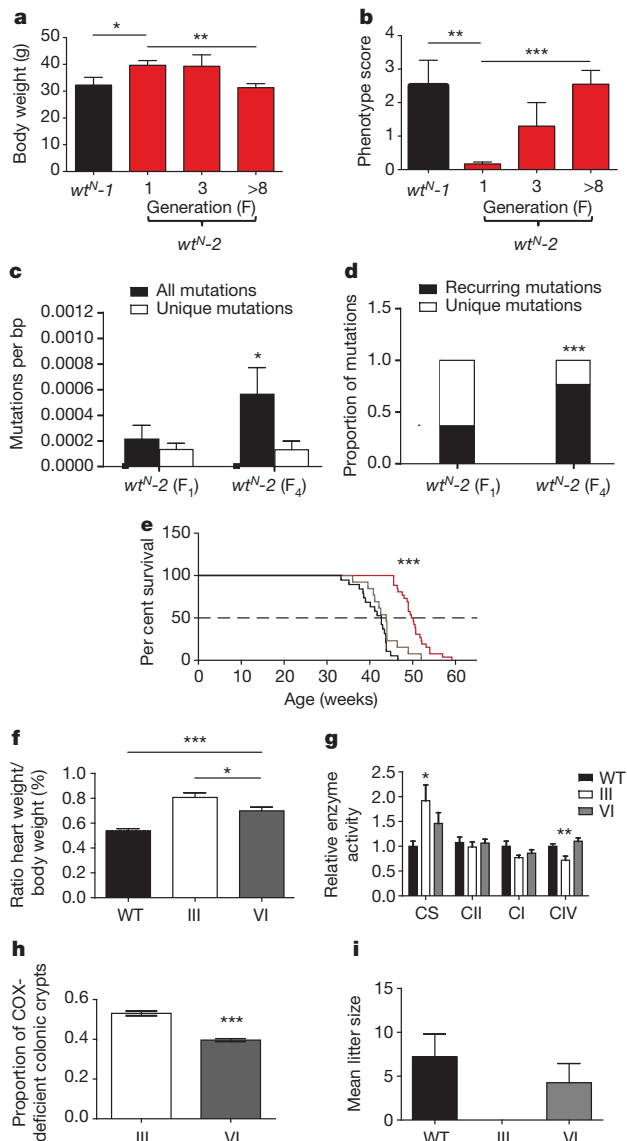
female (Fig. 2e, red line) and the proportion of  $PolgA^{mut/mut}$  mice born (Fig. 2f, red line). The reciprocal experiment wherein wild-type males were crossed with  $PolgA^{wt/mut}$  females to obtain  $PolgA^{wt/mut}$  animals for subsequent intercrossing did not modify the markedly impaired fecundity (Fig. 2d–f, grey lines). Ongoing mtDNA mutagenesis in the maternal germ line thus causes anticipation of fecundity phenotypes and this deterioration can be reversed by the introduction of wild-type mtDNA into females.

We proceeded to study the extent to which a burst of mutagenesis in the maternal germ line would result in mtDNA mutation transmission to subsequent generations. To this end, we bred  $PolgA^{wt/mut}$  females with re-introduced wild-type mtDNA ( $n = 8$ ) to wild-type males (Supplementary Fig. 2a) and followed segregation of mutated mtDNA in  $wt^N$  maternal lineages ( $n = 13$ ). Sequencing showed that heteroplasmic mtDNA mutations were present in all  $wt^N$  mice in backcross generations 2–8 ( $n = 387$  individuals). In total, 192 different clonally expanded mtDNA mutations were identified (Supplementary Fig. 2b).

Next, we generated three different types of  $wt^N$  mice:  $wt^N-1$  ( $wt^N$  mice obtained from a standard  $PolgA^{wt/mut}$  intercross (Fig. 1a)),  $wt^N-2$  ( $wt^N$  mice obtained from a  $PolgA^{wt/mut}$  intercross in which the female contains re-introduced mtDNA (Fig. 1b)) and  $wt^N-3$  ( $wt^N$  mice obtained from a  $PolgA^{wt/mut}$  intercross in which the male contains re-introduced mtDNA). The  $wt^N-1$  mice had significantly lower body weight than  $wt^N-2$  mice (Fig. 3a), whereas there was a similar reduction in body weight in  $wt^N-1$  and  $wt^N-3$  mice (Supplementary Fig. 3a). The body weights of  $wt^N-2$  mice decreased as crosses were continued without re-introduction of wild-type mtDNA (Fig. 3a). At 65 weeks of age,  $wt^N-1$  mice showed a significant ageing-associated phenotype (mean score = 2.5; s.d. = 1.8), whereas  $wt^N-2$  mice demonstrated virtually no such phenotypes (mean score = 0.16, s.d. = 0.2,  $P = 0.008$ ) (Fig. 3b). The score increased in successive generations of  $wt^N-2$  mice (Fig. 3b) and was increased to a similar extent in  $wt^N-1$  and  $wt^N-3$  mice (Supplementary Fig. 3b). We also analysed behaviour and found that spontaneous rearing was higher in  $wt^N-2$  than in  $wt^N-1$  mice ( $P = 0.03$ ) and that it decreased in successive generations of  $wt^N-2$  mice ( $P = 0.003$ ) (Supplementary Fig. 3c). There was no difference in rearing between  $wt^N-1$  and  $wt^N-3$  mice (Supplementary Fig. 3d), whereas we found impaired rearing in 2-year-old  $wt^N-1$  mice in comparison with 2.3-year-old C57BL/6N mice ( $P = 0.009$ ) (Supplementary Fig. 3e). These results prompted us to analyse mtDNA mutation levels and we found that the total mutation loads and the proportion of recurring mutations increased in  $wt^N-2$  animals between intercross generations  $F_1$  and  $F_4$  (Fig. 3c, d and Supplementary Table 1).

We observed that  $PolgA^{mut/mut}$  mice derived from the standard breeding had a mean lifespan of 42.7 weeks ( $n = 19$ ) (Fig. 3e, black line) as previously reported<sup>21</sup>, whereas  $PolgA^{mut/mut}$  mice, whose  $PolgA^{wt/mut}$  mothers lacked maternally transmitted mtDNA mutations, had a mean lifespan of 50.2 weeks ( $n = 26$ ) (Fig. 3e, red line). This increase in lifespan was statistically significant ( $P < 0.0001$ ) and was completely reversed in  $PolgA^{mut/mut}$  mice born to  $PolgA^{wt/mut}$  mothers maintained by intercrossing for several generations after re-introduction of wild-type mtDNA (mean lifespan ~43.5 weeks,  $n = 13$ ) (Fig. 3e, grey line).

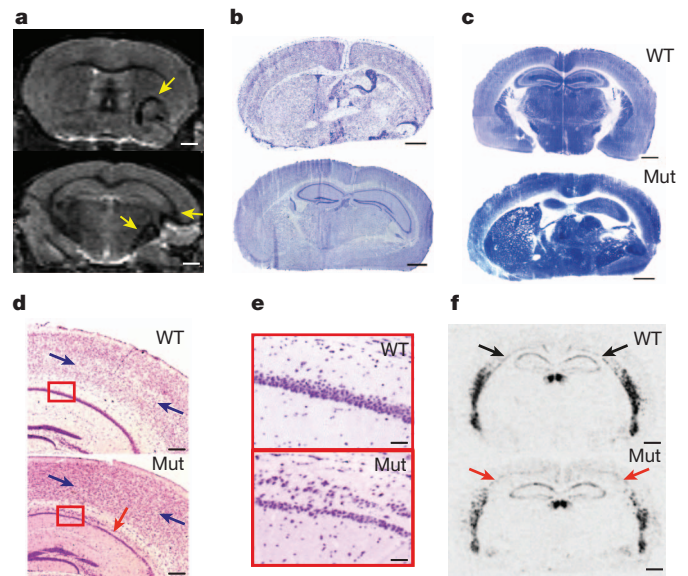
Detailed phenotyping was performed for mtDNA mutator mice of type III and VI at age ~35 weeks and revealed a comparable reduction in body mass (Supplementary Fig. 4a), the reduction in haemoglobin concentration (Supplementary Fig. 4b), the reduction in erythrocyte count (Supplementary Fig. 4c) and the degree of compensatory splenomegaly (Supplementary Fig. 4d) were almost identical, indicating a similar reduction of haematopoietic stem cell capacity in both types of mice. The ratio of heart weight to body weight was increased in both mouse types ( $P < 0.0001$ , analysis of variance (ANOVA)) (Fig. 3f), consistent with cardiomyopathy development<sup>21</sup>. However, the hearts from type III mice were more enlarged than those from type VI mice (Fig. 3f;  $P = 0.023$ , two-tailed  $t$ -test, Welch's correction). Consistently, the complex IV (cytochrome  $c$  oxidase (COX)) activity was reduced in type III, but not type VI, mice (Fig. 3g). Clonally expanded mtDNA mutations



**Figure 3 | Premature ageing phenotypes in mice with a wild-type nuclear genome and in mtDNA mutator mice.** **a**, Body weight at ~65 weeks in  $wt^{N-1}$  ( $n = 6$ ) and  $wt^{N-2}$  mice (F<sub>1</sub>,  $n = 13$ ; F<sub>3</sub>,  $n = 5$ ; F<sub>8</sub>,  $n = 10$ ). **b**, Phenotypic ageing score at ~65 weeks in  $wt^{N-1}$  and  $wt^{N-2}$  mice, Kruskal-Wallis or Mann-Whitney, post-hoc significances displayed. **c**, mtDNA mutation load comparisons of  $wt^{N-2}$  F<sub>1</sub> ( $n = 6$  animals) and F<sub>4</sub> ( $n = 3$  animals). **d**, Proportions of unique and repeated mtDNA mutations of  $wt^{N-2}$  animals from F<sub>1</sub> and F<sub>4</sub>. Levels of clonally expanded point mutations increase from F<sub>1</sub> to F<sub>4</sub>. **e**, Lifespan of  $PolgA^{wt/mut}$  mice obtained from standard intercrosses of  $PolgA^{wt/mut}$  mice (black line,  $n = 19$ ), intercrosses of  $PolgA^{wt/mut}$  mice after re-introduction of wild-type mtDNA into females (red line,  $n = 26$ ) and intercrosses of  $PolgA^{wt/mut}$  mice obtained after crossing wild-type males to  $PolgA^{wt/mut}$  females (grey line,  $n = 13$ ). Mantel-Cox test. **f**, The ratio of heart to body weight in 35-week-old wild-type mice (WT,  $n = 10$ ), type III mice ( $n = 9$ ) and type VI mice ( $n = 9$ ).  $t$ -tests with Welch's correction. **g**, Respiratory chain function in heart at 35 weeks of age from wild-type mice ( $n = 5$ ), type III mice ( $n = 5$ ) and type VI mice ( $n = 5$ ). Relative enzyme activities of citrate synthase (CS), succinate dehydrogenase (CII), NADH dehydrogenase (CI) and COX (CIV). Two-tailed, unpaired  $t$ -tests. **h**, Proportion of COX-deficient colonic crypts at 35 weeks in type III and type VI mice. No COX deficiency was present in age-matched controls ( $n = 5$  per genotype). **i**, Mean litter size after mating wild-type females to type III males ( $n = 8$  matings) or to type VI males ( $n = 8$  matings). The average litter size of wild-type mice is shown (black bar). One-way ANOVA (**a**, **b**), unpaired two-tailed  $t$ -test (**c**, **f**, **g**), Fisher's exact (**d**). \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ . Error bars denote  $\pm$  s.e.m. (**a**, **b**, **f**, **g**, **h**) and  $\pm$  s.d. (**c**, **i**).

typically create a mosaic pattern of respiratory-chain-deficient cells in the colons of mice and humans<sup>26,27</sup> and such COX-deficient colonic crypts were abundant in both type III and VI mice (Fig. 3h and Supplementary Fig. 4e). However, the proportion of COX-deficient colonic crypts was significantly higher in type III than type VI mice (Fig. 3h), indicating increased levels of clonally expanded mtDNA mutations. We assessed male fertility and found that type III mice were sterile (Fig. 3i), as previously reported<sup>21</sup>, whereas type VI males were fertile and could produce offspring, albeit with reduced litter sizes (mean =  $4.3 \pm 2.2$ ) in comparison to controls ( $7.2 \pm 2.6$  pups per litter) (Fig. 3i). These findings show that the ongoing massive somatic mutagenesis of mtDNA is sufficient on its own to drive the haematological stem cell phenotype<sup>28–30</sup>, whereas maternally transmitted mtDNA mutations contribute to phenotypes in other rapidly proliferating tissues, such as testis and colon, and in postmitotic tissues, such as heart, of mtDNA mutator mice<sup>21,22</sup>.

Unexpectedly, when performing magnetic resonance imaging (MRI) studies of ageing phenotypes, we found a gross brain malformation in an mtDNA mutator mouse (Fig. 4a), which, on subsequent histology, was shown to be due to mosaic subcortical and striatal perturbations, as well as malformations affecting both the right and left hippocampus (Fig. 4b). Succinate dehydrogenase enzyme histochemistry of the same brain revealed areas of increased activity (Fig. 4c), suggesting increased cell density and/or compensatory mitochondrial biogenesis. On the



**Figure 4 | Focal and symmetric brain malformations in mtDNA mutator mice.** **a**, MRI at two levels showing an abnormal signal from white matter in striatum (arrow, top; left arrow, bottom) and elongation of the right hippocampal profile (right arrow, bottom) in a 24-week-old male mtDNA mutator mouse. **b**, Cresyl violet staining of brain sections at two levels showing severe striatal perturbations (top) and malformation of hippocampus (bottom) in the same mouse as the one depicted in panel a. **c**, Succinate dehydrogenase staining of brain sections of a wild-type and the mtDNA mutator mouse (Mut) depicted in panel a. **d**, Cresyl violet staining showing cortical and hippocampal lamination disturbances, primarily affecting layers III–IV (blue arrows) in cerebral cortex and CA1 pyramidal cells in hippocampus (red arrow), in a 12-week-old female mtDNA mutator mouse (Mut). Analysis of an age-matched wild-type mouse is also shown (WT). **e**, Increased magnification of boxed areas in panel d, showing hippocampal lamination disturbances in the mtDNA mutator mouse. Notably, stratum pyramidale is separated into two layers, with hippocampal pyramidal cells localized both in an outer, thin, less orderly layer and an inner, thicker, more orderly layer (bottom). Analysis of an age-matched wild-type mouse is also shown (top). **f**, *In situ* hybridization to detect nuclear receptor related 1 (*Nurr1*) mRNA in the mtDNA mutator (Mut) mouse depicted in panel d. *Nurr1*-expressing neurons are found in layers III–IV of the mtDNA mutator mouse (red arrows, bottom), contrasting with the normal expression pattern in layers V–VI (black arrows, top). Scale bars, 1 mm (**a**–**c**, **f**), 500  $\mu$ m (**d**) and 125  $\mu$ m (**e**).



basis of this observation, we performed a systematic study of brain morphology in mtDNA mutator mice from the standard breeding (type III mice). Microscopic, and occasionally also macroscopic, brain malformations were observed in ~32% of the mtDNA mutator mice ( $n = 13$  of 41 mice), but not in wild-type mice ( $n = 0$  of 46 mice) at ages 7–46 weeks. Brain malformations were equally common in both sexes (41.5% female and 58.5% male) and most malformations were focal (69%) and manifested as different forms of brain asymmetries (Fig. 4a–c). Some mtDNA mutator mice ( $n = 4$ ) had widespread bilateral brain malformations, which included cortical and hippocampal lamination disturbances (Fig. 4d–f), hypertrophic/hyperplastic changes in hippocampus and cerebral cortex and hyperplasia in cerebellum. Brain malformations were not observed in type I or II mice, despite the fact that they contain maternally transmitted mtDNA mutations, or in mtDNA mutator mice ( $n = 11$ ) born to *PolgA*<sup>wt/mut</sup> mothers with re-introduced wild-type mtDNA. These findings strongly suggest that a dual hit—that is, the presence of maternally inherited mtDNA mutations in combination with massive somatic mutagenesis of mtDNA—is required for brain malformations to occur.

Our findings show that maternally transmitted mtDNA mutations provide a baseline mutation load upon which somatic mutagenesis will act. We show that low levels of germline-transmitted mtDNA mutations per se can have life-long consequences and cause premature ageing. We also show that germline-transmitted mtDNA mutations constitute a risk factor that may lead to developmental disturbances if combined with increased somatic mutagenesis of mtDNA. These findings suggest that inherited human mtDNA sequence variation and low-level heteroplasmy combined with somatic mtDNA mutagenesis may have an additive effect in creating phenotypes relevant for human pathology and ageing.

## METHODS SUMMARY

Mice expressing the exonuclease-deficient mtDNA polymerase gamma were maintained on the inbred C57BL/6NCRl background. The study was performed in strict accordance with guidelines of the Federation of the European Laboratory Animal Science Association. Protocols were approved by the Landesamt für Natur, Umwelt und Verbraucherschutz, Nordrhein-Westfalen, Germany, and the animal welfare ethics committee in Sweden. Complete mtDNA sequencing was carried out as described<sup>11</sup>, with a modified primer set (Supplementary Table 3). Mutation loads were analysed by the cloning of mtDNA fragments amplified with a high-fidelity polymerase, and individual clones were sequenced. Data were filtered against nuclear pseudogene sequences of mitochondrial origin. The data were confirmed by direct sequencing of  $\lambda$ -phage-cloned complete mtDNAs, and clonal expansion was confirmed by quantitative restriction fragment length polymorphism (RFLP)-PCR analyses. Blood analyses were performed by Laboklin GmbH & Co. KG. Clonal expansion analysis of mouse colonic tissues was performed as described<sup>26,27</sup>. Brain morphology was analysed by MRI with a BioSpec Avance 47/40 (Bruker) or by histochemistry of brain sections stained with either cresyl violet or haematoxylin and eosin.

**Full Methods** and any associated references are available in the online version of the paper.

**Received 8 January; accepted 16 July 2013.**

**Published online 21 August 2013.**

- Szilard, L. On the nature of the aging process. *Proc. Natl Acad. Sci. USA* **45**, 30–45 (1959).
- Kirkwood, T. B. L. Understanding the odd science of aging. *Cell* **120**, 437–447 (2005).
- Ernster, L., Löw, H., Nordenbrand, K. & Ernster, B. A component promoting oxidative phosphorylation, released from mitochondria during aging. *Exp. Cell Res.* **9**, 348–349 (1955).
- Miquel, J., Economos, A. C., Fleming, J. & Johnson, J. E. Mitochondrial role in cell aging. *Exp. Gerontol.* **15**, 575–591 (1980).
- Pikó, L., Hougham, A. J. & Bulpitt, K. J. Studies of sequence heterogeneity of mitochondrial DNA from rat and mouse tissues: evidence for an increased frequency of deletions/additions with aging. *Mech. Ageing Dev.* **43**, 279–293 (1988).
- Müller-Höcker, J. Cytochrome-c-oxidase deficient cardiomyocytes in the human heart — an age-related phenomenon. *Am. J. Pathol.* **134**, 1167–1173 (1989).
- Cortopassi, G. A. & Arnheim, N. Detection of a specific mitochondrial DNA deletion in tissues of older humans. *Nucleic Acids Res.* **18**, 6927–6933 (1990).
- Corral-Debrinski, M. et al. Mitochondrial DNA deletions in human brain: regional variability and increase with advanced age. *Nature Genet.* **2**, 324–329 (1992).
- Pakendorf, B. & Stoneking, M. Mitochondrial DNA and human evolution. *Annu. Rev. Genomics Hum. Genet.* **6**, 165–183 (2005).
- Krakauer, D. C. & Mira, A. Mitochondria and germ-cell death. *Nature* **400**, 125–126 (1999).
- Stewart, J. B. et al. Strong purifying selection in transmission of mammalian mitochondrial DNA. *PLoS Biol.* **6**, e10 (2008).
- Fan, W. et al. A mouse model of mitochondrial disease reveals germline selection against severe mtDNA mutations. *Science* **319**, 958–962 (2008).
- Freyer, C. et al. Variation in germline mtDNA heteroplasmy is determined prenatally but modified during subsequent transmission. *Nature Genet.* **44**, 1282–1285 (2012).
- Stewart, J. B., Freyer, C., Elson, J. L. & Larsson, N.-G. Purifying selection of mtDNA and its implications for understanding evolution and mitochondrial disease. *Nature Rev. Genet.* **9**, 657–662 (2008).
- Li, M. et al. Detecting heteroplasmy from high-throughput sequencing of complete human mitochondrial DNA genomes. *Am. J. Hum. Genet.* **87**, 237–249 (2010).
- Payne, B. A. I. et al. Universal heteroplasmy of human mitochondrial DNA. *Hum. Mol. Genet.* **22**, 384–390 (2013).
- Payne, B. A. I. et al. Mitochondrial aging is accelerated by anti-retroviral therapy through the clonal expansion of mtDNA mutations. *Nature Genet.* **43**, 806–810 (2011).
- Greaves, L. C. & Turnbull, D. M. Mitochondrial DNA mutations and ageing. *Biochim. Biophys. Acta* **1790**, 1015–1020 (2009).
- Larsson, N. G. Somatic mitochondrial DNA mutations in mammalian aging. *Annu. Rev. Biochem.* **79**, 683–706 (2010).
- Ameur, A. et al. Ultra-deep sequencing of mouse mitochondrial DNA: mutational patterns and their origins. *PLoS Genet.* **7**, e1002028 (2011).
- Trifunovic, A. et al. Premature ageing in mice expressing defective mitochondrial DNA polymerase. *Nature* **429**, 417–423 (2004).
- Kujth, G. C. et al. Mitochondrial DNA mutations, oxidative stress, and apoptosis in mammalian aging. *Science* **309**, 481–484 (2005).
- Hance, N., Ekstrand, M. I. & Trifunovic, A. Mitochondrial DNA polymerase gamma is essential for mammalian embryogenesis. *Hum. Mol. Genet.* **14**, 1775–1783 (2005).
- Kraytsberg, Y. & Khrapko, K. Single-molecule PCR: an artifact-free PCR approach for the analysis of somatic mutations. *Expert Rev. Mol. Diagn.* **5**, 809–815 (2005).
- Greaves, L. C. et al. Quantification of mitochondrial DNA mutation load. *Ageing Cell* **8**, 566–572 (2009).
- Greaves, L. C., Elson, J. L., Nooteboom, M. & Grady, J. P. Comparison of mitochondrial mutation spectra in ageing human colonic epithelium and disease: absence of evidence for purifying selection in somatic mitochondrial DNA point mutations. *PLoS Genet.* **8**, e1003082 (2012).
- Taylor, R. W. et al. Mitochondrial DNA mutations in human colonic crypt stem cells. *J. Clin. Invest.* **112**, 1351–1360 (2003).
- Ahlqvist, K. J. et al. Somatic progenitor cell vulnerability to mitochondrial DNA mutagenesis underlies progeroid phenotypes in *Polg* mutator mice. *Cell Metab.* **15**, 100–109 (2012).
- Norddahl, G. L. et al. Accumulating mitochondrial DNA mutations drive premature hematopoietic aging phenotypes distinct from physiological stem cell aging. *Stem Cells* **8**, 499–510 (2011).
- Chen, M. L. et al. Erythroid dysplasia, megaloblastic anemia, and impaired lymphopoiesis arising from mitochondrial dysfunction. *Blood* **114**, 4045–4053 (2009).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** The study was supported by ERC Advanced Investigator grants (268897 to N.-G.L. and 322744 to L.O.), the Swedish Research Council (K2011-62X-21870-01-6 to N.-G.L. and K2012-62X-03185-42-4 to L.O.), the Swedish Brain Foundation (N.-G.L. and J.M.R.), Swedish Brain Power (L.O. and J.M.R.), the Swedish Parkinson Foundation (N.-G.L.), the Karolinska Distinguished Professor Award (L.O.), the Swedish Alzheimer Foundation (L.O.), the National Institutes of Health (AG04418 to L.O. and NS070825 to B.J.H.), the National Institute on Drug Abuse (J.M.R.), the National Institutes of Health/Karolinska Institutet Graduate Partnerships Program (J.M.R.) and the Swedish Society for Medical Research (G.C.). J.B.S. acknowledges support from the United Mitochondrial Disease Foundation.

**Author Contributions** J.M.R., J.B.S. and G.C. performed breeding and phenotypic analyses of mice. J.B.S., E.H. and C.F. performed mtDNA sequence analysis. J.M.R., J.B.S. and S.B. performed histology and MRI analyses. A.M. and M.L. performed molecular analyses and measurement of respiratory chain function. J.M.R., J.B.S., B.J.H., L.O. and N.-G.L. conceived the ideas, designed the experiments and wrote the paper.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to L.O. ([lars.olson@ki.se](mailto:lars.olson@ki.se)) or N.-G.L. ([larsson@age.mpg.de](mailto:larsson@age.mpg.de)).

## METHODS

**Animals.** All mice in this study were on an inbred C57BL/6NCrl nuclear background. The study was performed in strict accordance with the recommendations and guideline of the Federation of the European Laboratory Animal Science Association (FELASA). Protocols were approved by the Landesamt für Natur, Umwelt und Verbraucherschutz, Nordrhein-Westfalen, Germany, and, in Sweden, the animal welfare ethics committee and performed in accordance with Swedish law.

**Complete mtDNA sequencing.** The protocol used was modified slightly from that described previously<sup>11</sup>. Total DNA from tail biopsies from 3-week-old animals was collected, and the mitochondrial genome was amplified with a modified primer set, producing 30 overlapping fragments of 1–1.5 kilobases (kb) (Supplementary Table 3). Each primer was tagged with M13F or M13R sequence for use in sequencing reactions<sup>11</sup>. Sequences were aligned to the C57BL/6NCrl mtDNA reference sequence (JF286601.1) to detect mutations, with position numbering altered to remain consistent with the C57BL/6J mtDNA reference sequence (NC\_005089.1). Only mutations >20% in relative abundance are detected by this assay<sup>31</sup>, thus we excluded animals from the first backcross generation (maximum mutation levels 15%) from this assay.

**Mutation load analysis.** A PCR–clone–sequence method was used to assay mutation loads. Total DNA was extracted from tail biopsy samples obtained from the mice at weaning (~3 weeks of age). Detailed protocols have been published elsewhere<sup>32</sup>. Two pseudogenes of the target region are found in the nuclear genome, but differ from the reference sequence by 19–20 sites. Any sequence sharing two or more consecutive single nucleotide polymorphisms with these pseudogenes was excluded from the analysis. An error-rate assessment of the method is described in ref. 32. One single nucleotide polymorphism variation was observed in 295 clones assayed, giving an observed error rate of  $3.48 \times 10^{-6}$ . This is less than one-sixth the mean rate observed for the wild-type C57BL/6N. We obtained 3.3 megabases of mtDNA sequence (mean coverage = 88.7 kb per animal from an average of 91 clones per animal) from type I–VI mice (Fig. 1a–c).

Lambda clones of intact mtDNA molecules were obtained as described<sup>20</sup>. Mitochondria were sucrose purified from livers of 10-week-old animals to obtain sufficient mtDNA for cloning. Approximately 30 clones were chosen at random for complete mtDNA sequencing, as described above. In total, 250,298 bp and 259,858 bp of cloned mtDNA was sequenced from two C57BL/6N controls and showed no mtDNA mutations (error rate  $<3.85 \times 10^{-6}$  mutations per bp at 16× sequencing depth). In total, 350 complete mtDNA molecules, corresponding to a total of 5.6 megabases of mtDNA, were sequenced.

**RFLP-PCR assay.** Mutation level quantification by quantitative restriction fragment length polymorphism (RFLP)-PCR was carried out as described<sup>13</sup>. One microlitre of lysate was used for PCR. Primers used: 5247A>G, 5'-TGTAAGAC GACGGCCAGTAGACCTCAACTAGATTGGCAG-3' and 5'-CAGGAAACAG CTATGACCGGTGACTGGCTGAGTAAGCATTAGACTG-3'; 5839T>C and 5840G>A, 5'-CCCACCTCTAGCCGGAATCTAGC-3' and 5'-CAGGAAACA GCTATGACCAATGCCTGCGGCTAGCACTG-3'. M13 tag sequences are in bold. FAM-labelled primers M13F (for 5247A>G) and M13R (for 5839T>C and 5840G>A) were used for fragment detection. Restriction enzymes used in the assay were HphI for 5247A>G, NlaIII for 5839T>C and 5840G>A (New England Biolabs).

**Statistical analysis.** Statistical analyses were performed using GraphPad Prism (GraphPad Software), with  $\alpha = 0.05$ .

**Mitochondrial respiratory chain complex activities.** Analysis of isolated respiratory chain complex activities was assayed as described<sup>13</sup>.

**Histology and enzyme histochemistry.** For visualization of brain morphology, slides were stained with either cresyl violet or haematoxylin and eosin, dehydrated

in increasing concentrations of ethanol (70%, 95% and 99.5%), mounted (Entellan, VWR International AB) and coverslipped. The activity of succinate dehydrogenase was visualized as previously described<sup>33</sup>.

**Tissue preparation for cryosectioning.** Fresh–frozen brain sections for histochemistry and *in situ* hybridization were prepared as described<sup>33</sup>. Combined COX and succinate dehydrogenase staining of mouse colons was conducted as described<sup>26</sup>. Counts of colonic crypts with and without COX activity were performed by a researcher blinded to animal genotype.

**Nurr1 gene expression.** High-stringency radioactive *in situ* hybridization was performed as previously described<sup>33</sup> to detect *Nurr1* mRNA transcripts using a synthetic DNA oligonucleotide probe (5'-GCGTAGTGGCCACGTAGTCTGG TGGAAAGTTCTGAAGGGAGCCCGGATCG-3'; Thermo Scientific).

**Magnetic resonance imaging.** MRI was performed on animals as described previously<sup>33</sup> using a horizontal 4.7 T/40 cm magnet (BioSpec Avance 47/40; Bruker), fitted with a 12-cm inner diameter self-shielded gradient system (maximum gradient strength 200 mT m<sup>-1</sup>). A linear birdcage resonator (25 mm inner diameter) was used for excitation and detection. The sequence used for inversion recovery imaging had the following parameters: echo time 35.6 ms, repetition time 2566.8 ms, rapid acquisition with relaxation enhancement (RARE)-factor 8 with RARE-maximum 4, inversion time 450 ms, and matrix size 64 × 64 × 128 with two averages. The field of view was 0.9 × 1.2 (or 1.4 for larger voxel) × 1.8 cm, yielding the following directional resolutions: 0.14 mm (dorsolateral), 0.18 mm (left–right) and 0.14 mm (rostrocaudal).

**Male fertility analysis.** Type III males ( $n = 8$ ) and type VI males ( $n = 8$ ) (Fig. 1a, c) were mated to single C57BL/6N females. The size of the first litter was recorded. Males were removed from females after 60 days of unproductive mating, with females observed for three additional weeks to confirm that the mating was unproductive.

**Blood and tissue analyses.** Blood was drawn from the heart immediately after death, with haemoglobin and erythrocyte values determined by Laboklin GmbH & Co. KG. Spleen and heart masses were determined after dissection and washing in PBS to remove residual blood.

**Rearing behaviour.** Spontaneous rearing activity in an open field was studied in wild-type mice derived from various crossings. A multi-cage infrared-sensitive motion-detection system (Accuscan Instruments) was used. Total rearing was recorded at 5-min intervals over 90 min, and was determined by vertical sensor beam-breaks using appropriate software (VersaMax activity monitor, Accuscan Instruments). Animals were habituated in a dimly lit, low-noise and ventilated experimental room kept at 20–22 °C for 1 h before testing. Experiments were performed between 11 am and 6 pm (light phase).

**Phenotype scoring.** The presence of alopecia, greying of the hair, body size reduction and kyphosis were each scored on a scale of 0–3 (0, absence of phenotype; 1, moderate presence of phenotype; 2, strong presence of phenotype; 3, severe presence of phenotype) and the scores were summed for each animal. A summed score of 0 thus indicates normal ageing, whereas a score of 12 indicates marked premature ageing similar to the end-stage phenotype of mtDNA mutator mice. The animals were scored blind without knowledge of the genotype.

- Hancock, D. K., Tully, L. A. & Levin, B. C. A Standard Reference Material to determine the sensitivity of techniques for detecting low-frequency mutations, SNPs, and heteroplasmies in mitochondrial DNA. *Genomics* **86**, 446–461 (2005).
- Wanrooij, S. *et al.* *In vivo* mutagenesis reveals that OriL is essential for mitochondrial DNA replication. *EMBO Rep.* **13**, 1130–1137 (2012).
- Ross, J. M. *et al.* High brain lactate is a hallmark of aging and caused by a shift in the lactate dehydrogenase A/B ratio. *Proc. Natl Acad. Sci. USA* **107**, 20087–20092 (2010).

# DNA damage in germ cells induces an innate immune response that triggers systemic stress resistance

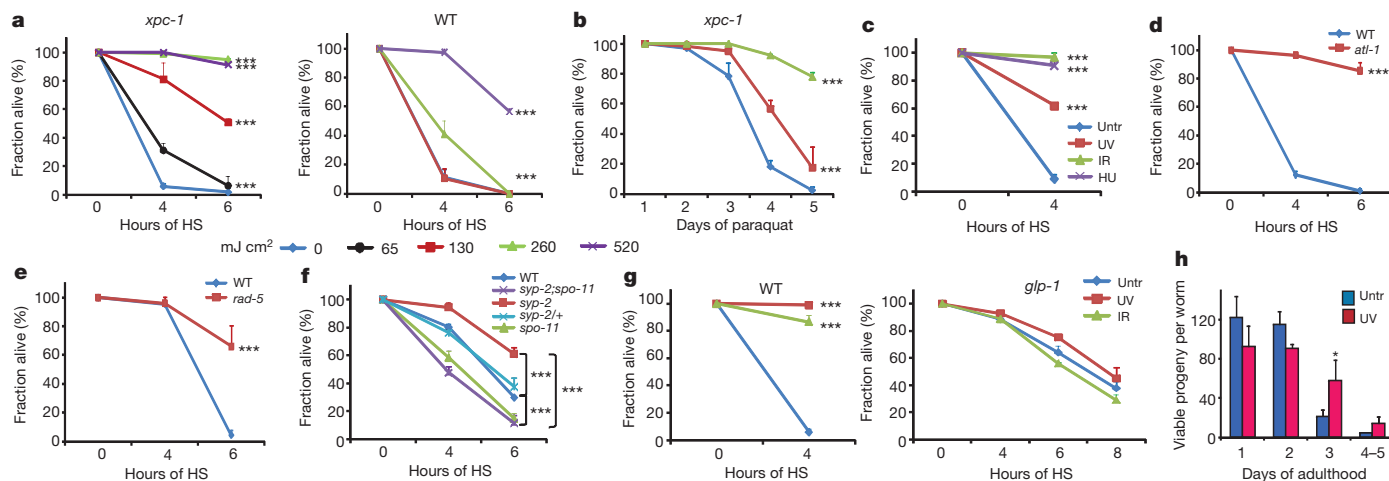
Maria A. Ermolaeva<sup>1</sup>, Alexandra Segref<sup>1</sup>, Alexander Dakhovnik<sup>1</sup>, Hui-Ling Ou<sup>1</sup>, Jennifer I. Schneider<sup>1</sup>, Olaf Utermöhlen<sup>2,3</sup>, Thorsten Hoppe<sup>1</sup> & Björn Schumacher<sup>1,4</sup>

DNA damage responses have been well characterized with regard to their cell-autonomous checkpoint functions leading to cell cycle arrest, senescence and apoptosis<sup>1</sup>. In contrast, systemic responses to tissue-specific genome instability remain poorly understood. In adult *Caenorhabditis elegans* worms germ cells undergo mitotic and meiotic cell divisions, whereas somatic tissues are entirely post-mitotic. Consequently, DNA damage checkpoints function specifically in the germ line<sup>2</sup>, whereas somatic tissues in adult *C. elegans* are highly radio-resistant<sup>3</sup>. Some DNA repair systems such as global-genome nucleotide excision repair (GG-NER) remove lesions specifically in germ cells<sup>4</sup>. Here we investigated how genome instability in germ cells affects somatic tissues in *C. elegans*. We show that exogenous and endogenous DNA damage in germ cells evokes elevated resistance to heat and oxidative stress. The somatic stress resistance is mediated by the ERK MAP kinase MPK-1 in germ cells that triggers the induction of putative secreted peptides associated with innate immunity. The innate immune response leads to activation of the ubiquitin–proteasome system (UPS) in somatic tissues, which confers enhanced proteostasis and systemic stress resistance. We propose that elevated systemic stress resistance promotes endurance of somatic tissues to allow delay of progeny production when germ cells are genomically compromised.

GG-NER-defective *xpc-1(tm3886-allele)* mutants and wild-type animals that were treated with low doses of ultraviolet-B (UVB) radiation

showed extended survival after heat and oxidative stress (Fig. 1a, b, and Supplementary Fig. 1) that are well known to impair somatic tissue functioning causing premature death<sup>5</sup>. Similarly, ionizing radiation and hydroxyurea treatment evoked heat stress resistance in adult worms (Fig. 1c). ATR-checkpoint-defective *atl-1(tm853)* and *clk-2(mn159)* (also known as *rad-5*) mutant strains, both of which carry endogenous DNA damage in germ cells<sup>6,7</sup>, were highly resistant to heat stress (Fig. 1d, e). We next tested whether meiotic double strand breaks (DSBs), which are induced by the SPO-11 endonuclease and persist into late pachytene stages in synaptonemal complex mutants *syp-2(ok307)* (ref. 8) may trigger somatic stress resistance. Indeed, *syp-2(ok307)* mutants show elevated heat stress resistance that is reverted in the absence of SPO-11 (Fig. 1f). In fact, *spo-11(me44)* mutant worms are more heat sensitive than wild-type animals, suggesting that even physiologically induced meiotic DSBs are sufficient to confer somatic stress resistance (Fig. 1f). To validate the requirement of the germ line for somatic stress resistance upon DNA damage, we tested *glp-1(e2141)* mutants that fail to develop a germ line. Indeed, germline-less *glp-1(e2141)* mutants did not show elevated heat stress resistance upon ultraviolet or ionizing radiation treatment (Fig. 1g and Supplementary Fig. 2). Together, these results indicate that DNA damage in germ cells causes somatic stress resistance.

Germline-depleted worms, such as *glp-1(e2141)* mutants, are known to display a constitutively high stress resistance through activation of the



**Figure 1 | DNA damage in the germ line leads to somatic stress resistance.** **a, b,** Worms were UVB-treated at the L4 stage and 48 h later exposed to heat shock (HS) at 35 °C (**a**) or 5mM paraquat (**b**). **c,** Wild-type (WT) worms were treated at L4 stage with UVB (520 mJ cm<sup>-2</sup>), ionizing radiation (IR) (90 gray (Gy)) or hydroxyurea (25 mM) and exposed to heat shock as in **a**. **d–f** Heat shock at 35 °C was applied on the second day of adulthood. **g,** Wild-type and *glp-1* mutants were treated as described in **a**. **h,** L4 worms were treated with

UVB (520 mJ cm<sup>-2</sup>) and live progeny were monitored from the time of exposure until the indicated days. Total viable offspring per worm: untreated 263 (s.d. ± 12); ultraviolet-treated, 256 (s.d. ± 6). Graphs show an average of three biological replicates for each experimental condition; representative data from one of three independent experiments. Error bars, s.d. For *n* values see Supplementary Table 3. \**P* < 0.05; \*\*\**P* < 0.0001; log rank analysis (**a–g**) and two-tailed *t*-test (**h**).

<sup>1</sup>Cologne Excellence Cluster for Cellular Stress Responses in Aging-Associated Diseases (CECAD), Institute for Genetics, University of Cologne, Zùlpicher Strasse 47a, 50674 Cologne, Germany. <sup>2</sup>Institute for Medical Microbiology, Immunology and Hygiene, Medical Center, University of Cologne, Goldenfelsstrasse 19-21, 50935 Cologne, Germany. <sup>3</sup>Center for Molecular Medicine Cologne (CMCC), University of Cologne, 50931 Cologne, Germany. <sup>4</sup>Systems Biology of Ageing Cologne, University of Cologne, 50937 Cologne, Germany.



DAF-16 and DAF-12 transcription factors<sup>9,10</sup>. However, *daf-16(mu86)* and *daf-12(rh61 rh411);daf-16(mgDf50)* double mutants showed ultraviolet- and ionizing-radiation-induced heat stress resistance, indicating that germline DNA-damage-mediated stress resistance is exerted independently of the established 'germline ablation pathway' (Supplementary Fig. 3).

We wondered whether the elevated stress resistance may support somatic endurance to influence offspring generation after DNA damage in germ cells. Although ultraviolet treatment reduced initial offspring generation, progeny production increased from the third day of adulthood, resulting in a similar number of total offspring (Fig. 1h). Conceptually similar to cellular DNA damage checkpoints that allow time for DNA repair, the somatic stress resistance might act as systemic DNA damage checkpoint that preserves somatic functions when offspring generation is delayed as a consequence of germ cell DNA damage.

To examine which DNA damage response mechanisms mediate systemic stress resistance, we tested the DNA damage checkpoint regulators HUS-1 and ATM-1, as well as the apoptosis-inducing CEP-1 (*C. elegans* p53-like protein 1) and executing caspase CED-3, all of which were dispensable for DNA-damage-induced heat stress resistance (Supplementary Fig. 4). The mitogen-activated protein kinases (MAPKs) JNK, p38, and ERK1 and ERK2 are also activated upon DNA damage in several species<sup>11–13</sup>. Both JNK-defective *jnk-1(gk7)*, and p38-defective *pmk-1(km25)* mutants showed elevated heat stress resistance upon genotoxic treatment. In contrast, ERK-defective *mpk-1(oz140)* mutants as well as worms with RNA interference (RNAi)-mediated knockdown of *mpk-1* failed to exert enhanced heat stress resistance upon UV or ionizing radiation treatment (Fig. 2a and Supplementary Fig. 5a). As both *mpk-1(oz140)* and *pmk-1(km25)* mutants displayed elevated baseline heat stress resistance, we used *pmk-1(RNAi)* to revert

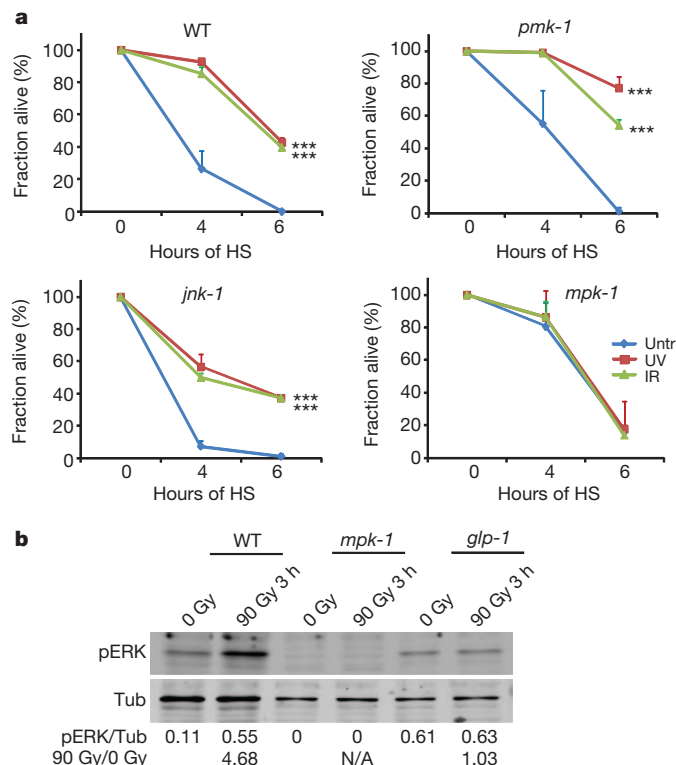
the compensatory baseline resistance of the *mpk-1(oz140)* mutant strain. Also *mpk-1(oz140)* mutants with *pmk-1(RNAi)* failed to evoke heat stress resistance after DNA damage (Supplementary Fig. 5b). MPK-1 that is activated upon ionizing radiation in meiotic pachytene cells<sup>11</sup> was phosphorylated upon ionizing radiation treatment in wild-type but not *glp-1(e2141)* mutants (Fig. 2b). Together, these results suggest that DNA damage-induced MPK-1 activation in germ cells instigates the systemic stress resistance.

To identify mediators of the systemic stress resistance that are induced by distinct types of DNA damage, we analysed the transcriptomes upon ionizing radiation<sup>14</sup> and UV treatment. Ionizing radiation treatment in wild-type and ultraviolet treatment in *xpc-1(tm3886)* mutant worms led to significant induction of 51 genes (Supplementary Table 1), among which the pathogen response C-type lectin protein domain<sup>15</sup> was overrepresented ( $P < 0.001$ ). Consistent with this, MPK-1 signalling has been shown to function in the response of *C. elegans* to *Microbacterium nematophilum* infection<sup>16</sup>. To assess whether the DNA damage response overlaps with the innate immune response we conducted comparative correlation analyses between the transcriptomes after DNA damage and infection with *M. nematophilum*<sup>17</sup> and *Pseudomonas aeruginosa* as well as *pmk-1* inactivation<sup>18</sup>. PMK-1 is required to induce the innate immune response to a variety of pathogens including *P. aeruginosa*<sup>19</sup>, and MPK-1- and PMK-1-mediated defence responses are significantly correlated, which is indicative of a common downstream effector gene expression program (Supplementary Fig. 6). Notably, gene expression changes after ultraviolet and ionizing radiation treatment ( $P < 0.01$ ) correlate with pathogen infection transcriptomes (Fig. 3a), suggesting similarities between DNA damage and innate immune responses.

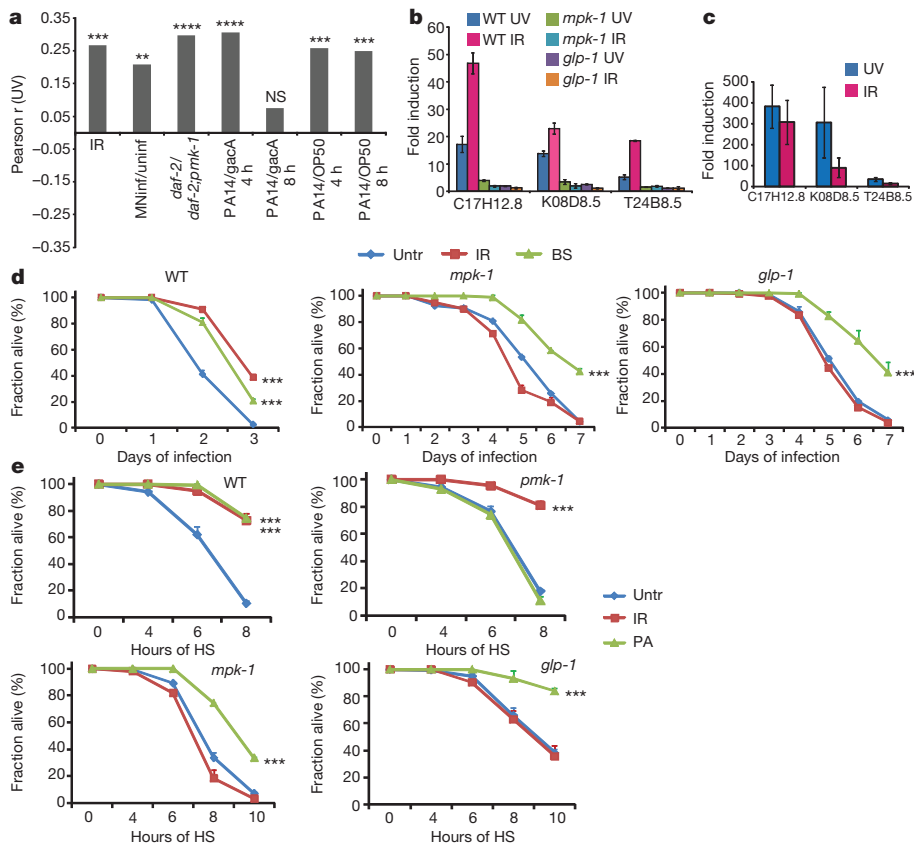
The nematode's innate immune response is mediated through the secretion of peptides that can be recognized by the presence of secretory signal peptides<sup>18,20</sup>. Out of the 51 ultraviolet- and ionizing radiation-induced genes, 35 carry signal peptides, and 33 of these are predicted to localize extracellularly (Supplementary Table 1). Among the 26 *pmk-1* and 34 *M. nematophilum* response genes that were also induced by ultraviolet and ionizing radiation treatment, 24 and 15 genes carried signal peptides, with 22 and 10 associated with extracellular secretion, respectively (Supplementary Table 2). Although wild-type worms induced the established putative secreted immune factors C17H12.8, K08D8.5, and T24B8.5 (refs 18, 21), *glp-1(e2141)* and *mpk-1(oz140)* mutants failed to induce their expression following ultraviolet or ionizing radiation treatment (Fig. 3b). Moreover, the expression of the immune factors was induced strongly in germ lines isolated from ultraviolet- and ionizing radiation-treated animals (Fig. 3c and Supplementary Fig. 7). These data suggest that DNA-damage-induced MPK-1 activation in the germ line leads to the induction of putative secreted immune peptides.

To evaluate whether DNA damage leads to a functional immune response, we tested survival after *P. aeruginosa* infection. Preconditioning with ionizing radiation or immunogenic—but non-pathogenic—*Bacillus subtilis* (Supplementary Fig. 8) led to enhanced survival in wild-type worms (Fig. 3d). *B. subtilis* but not ionizing radiation treatment led to elevated pathogen resistance in *mpk-1(oz140)* and *glp-1(e2141)* mutant animals (Fig. 3d), while *pmk-1(km25)* mutants developed enhanced pathogen resistance after ionizing radiation but not *B. subtilis* preconditioning (Supplementary Fig. 9). Consistent with a pathogen defence response, the intestinal *agl219* transgenic reporter that expresses green fluorescent protein (GFP) under the control of the T24B8.5 promoter<sup>22</sup> showed marked induction starting at 4 h after ultraviolet or ionizing radiation treatment in wild-type but not *glp-1(RNAi)* worms (Supplementary Fig. 10). We conclude that DNA damage-induced MPK-1 signalling in germ cells evokes a functional pathogen defence response.

We next examined whether the innate immune response after pathogen infection would provoke somatic stress resistance. Notably, pathogenic *P. aeruginosa* preconditioning resulted in elevated heat stress resistance (Fig. 3e). In contrast, *pmk-1(km25)* mutant worms



**Figure 2 | Stress resistance induced by germline DNA damage is mediated through MPK-1.** **a**, L4 larvae were exposed to 520 mJ cm<sup>-2</sup> UVB or 90 Gy IR and subjected to heat stress as described in Fig. 1. Sample collection as in Fig. 1. **b**, Worms were kept at 25 °C and IR-treated on day 1 of adulthood. Phosphorylated ERK (pERK) signal was compared to tubulin. Representative data from one of three independent experiments. Error bars, s.d. for *n* values see Supplementary Table 3; Tub, tubulin; \*\*\* $P < 0.0001$ , log rank test.



**Figure 3 | Somatic stress resistance after DNA damage in germ cells is mediated through MPK-1-induced functional innate immune response.**

**a**, Pearson correlation analysis of genes that were significantly induced after ultraviolet (UV) treatment with transcriptomes upon IR, pathogen infections. MN, *M. nematophilum*; PA, *P. aeruginosa* and *mpk-1* mutants. \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$ , two-tailed  $t$ -test.

**b**, Gene expression was assayed by quantitative polymerase chain reaction (qPCR) 6 h post UVB and IR treatment. **c**, Germ lines were dissected 1–2 h post UVB (520 mJ cm<sup>-2</sup>) or IR (90 Gy) for qPCR analysis. Panels **b** and **c** show representative data from one of three independent experiments.

**d**, L4 larvae were treated with 90 Gy IR or placed for 10 h on *B. subtilis* (BS), and exposed to *P. aeruginosa* 24 h later. **e**, Young adults were treated with 90 Gy IR or *P. aeruginosa* for 4 h, and were exposed to heat stress 24 h later. Error bars, s.d. Sample collection in panels **d** and **e** as in Fig. 1; for  $n$  values see Supplementary Table 3; NS, not significant; \*\*\* $P < 0.0001$ , log rank test.

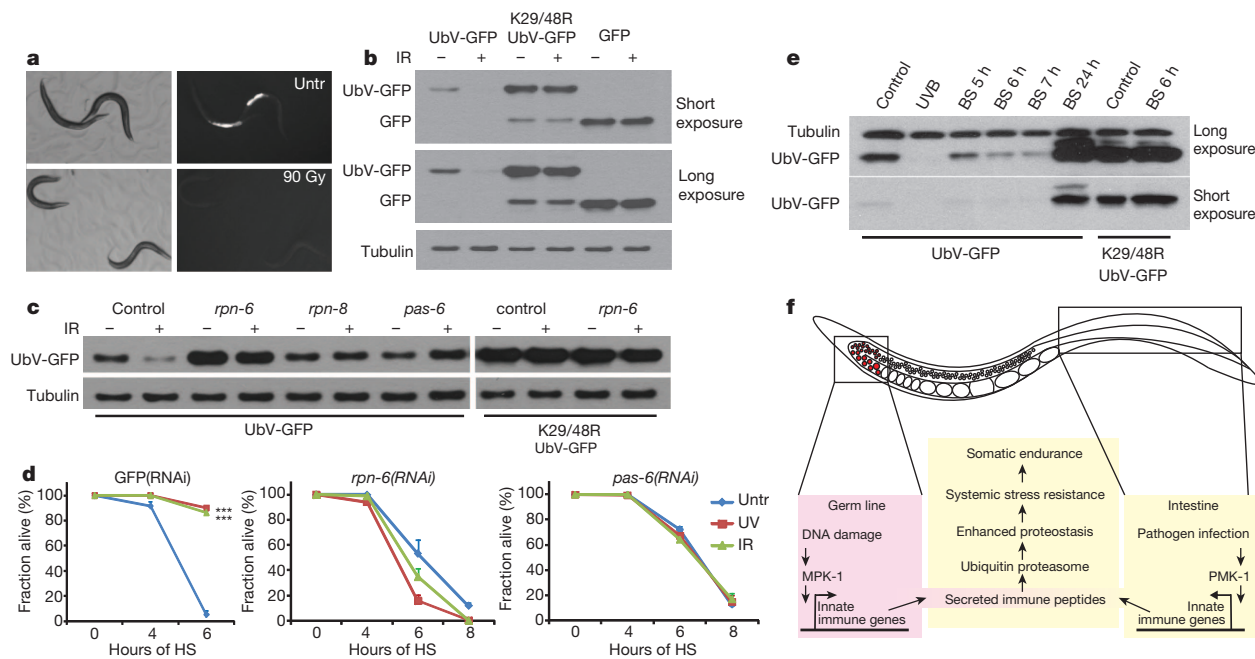
failed to develop enhanced heat stress resistance in response to infection, although they were proficient for ionizing radiation-induced heat stress resistance (Fig. 3e). Conversely, *mpk-1*(*oz140*) and *glp-1*(*e2141*) mutants that failed to induce heat stress resistance upon ionizing radiation treatment did show heat stress resistance upon *P. aeruginosa* infection (Fig. 3e and Supplementary Fig. 11). To address whether the innate immune response was required for inducing somatic stress resistance upon DNA damage we used *atf-7*(*qd22*) gain-of-function mutants that constitutively repress the transcription of innate immune genes<sup>21</sup> (Supplementary Fig. 12a). Ultraviolet- and ionizing radiation-induced heat stress resistance was strongly reduced in *atf-7*(*qd22*) mutants and reconstituted in intragenic revertant mutants *atf-7*(*qd22 qd130*) (Supplementary Fig. 12b). Together, these results indicate that the innate immune response is necessary and sufficient to mediate heat stress resistance after germline DNA damage.

To address the effector pathway in somatic tissues, we assessed HSF-1-mediated chaperone induction by following the expression of the heat shock protein *hsp-70* promoter fused to GFP<sup>5</sup>. *hsp-70::GFP* was not induced after ultraviolet or ionizing radiation treatment, and HSF-1 was largely dispensable for DNA-damage-induced somatic stress resistance (Supplementary Fig. 13). To assess altered protein turnover in somatic tissues we used a *sur-5::UbV-GFP* reporter strain that expresses ubiquitin with an C-terminal glycine to valine mutation fused to GFP under the control of the *sur-5* promoter in various somatic tissues and allows monitoring of UPS activity<sup>23</sup>. Notably, protein levels of UbV-GFP were reduced strongly after ionizing radiation or ultraviolet exposure, whereas both non-degradable UbV<sup>K29/48R</sup>-GFP or GFP alone, and messenger RNA levels were not altered (Fig. 4a, b and Supplementary Fig. 14), indicating enhanced ubiquitin-mediated protein turnover. We next assessed whether UPS activity was required for germline DNA-damage-induced stress resistance. Indeed, depletion of the 19S cap structure components *rpn-6*, *rpn-8*, or the 20S core particle component *pas-6* resulted in failure to degrade UbV-GFP and to evoke heat stress resistance upon genotoxic insult (Fig. 4c, d). The immune

reporter induction was not altered by *rpn-6* knockdown, suggesting that the immune response indeed functions upstream of UPS activation (Supplementary Fig. 15). In contrast to non-degradable UbV<sup>K29/48R</sup>-GFP, transient immunogenic *B. subtilis* exposure strongly decreased UbV-GFP protein levels (Fig. 4e). However, constitutive feeding with *B. subtilis* resulted in accumulation of UbV-GFP indicative of attenuated UPS activity after long immunogenic exposure (Fig. 4e). Together, these data establish that the transient activation of the innate immune response triggers UPS activation, which in turn promotes elevated systemic stress resistance.

We have uncovered a systemic response to DNA damage in germ cells that confers protection of somatic tissues against multiple stress factors and that we propose to name ‘germline DNA-damage-induced systemic stress resistance’ (GDISR) (Fig. 4f). The GDISR is triggered by MPK-1 activity in germ cells, whereas intestinal pathogen infection activates stress resistance through PMK-1. Both MAPKs induce putative secreted immune peptides that we suggest mediate systemic stress resistance by activation of the UPS. We propose that UPS activity in turn enhances proteostasis to elevate somatic endurance in the presence of genomically compromised germ cells or pathogen infection. Immune reactions to DNA damage also occur in mammals. In human cells the inflammasome can recognize cytosolic DNA<sup>24</sup> and genotoxic stress leads to induction of natural killer cell receptor NKG2D (also known as KLRK1) ligands<sup>25</sup> and cytokine secretion<sup>26</sup>. In addition to causing inflammation, the mammalian innate immune system has recently been implicated in the maintenance of tissue homeostasis and regeneration<sup>27,28</sup>. It will be very interesting to determine whether innate immune responses to DNA damage in humans might systemically enhance tissue maintenance before chronic inflammation is manifested.

As the recognition of DNA lesions is highly sensitive with a low threshold for triggering DNA damage responses, activation of the UPS may ameliorate protein turnover already at low damage levels before massive protein misfolding ensues in a toxic environment. In addition, the UPS



**Figure 4 | Innate immune responses trigger UPS activation to confer systemic stress resistance.** **a**, **b**, *sur-5::UbV-GFP* reporter expression was assessed 24 h post treatment with 90 Gy IR (**a**, bottom two panels) at L4 stage. Untr, untreated (**a**, top two panels). **c**, Early L4 *sur-5::UbV-GFP* reporter worms and *sur-5::UbV<sup>K29/48R</sup>-GFP* transgenic worms were fed for 12 h with RNAi against proteasomal subunits and then treated and analysed as in **b**. OP50 bacteria served as control. **d**, Worms were fed RNAi as in **c**, treated with UVB

(520 mJ cm<sup>-2</sup>) or IR (90 Gy) and exposed to heat stress as described in Fig. 1. Sample collection as in Fig. 1; for *n* values see Supplementary Table 3. Error bars, s.d.; \*\*\**P* < 0.0001, log rank analysis. **e**, *sur-5::UbV-GFP* and *sur-5::UbV<sup>K29/48R</sup>-GFP* reporter worms were treated with 520 mJ cm<sup>-2</sup> UVB or fed with *B. subtilis* at L4 stage for the indicated times, and protein levels were analysed on day 1 of adulthood. Panels **a–c** and **e** show representative data from one of three independent experiments. **f**, Model depicting the GDISR.

may facilitate the innate immune response by alleviating pressure from the protein folding machinery similar to the unfolded protein response<sup>29</sup>. We propose that GDISR comprises an ancestral somatic stress resistance program that prolongs somatic preservation through UPS-mediated enhanced proteostasis when genomically compromised germ cells require extended somatic endurance to ensure offspring generation.

## METHODS SUMMARY

Worms were grown under standard laboratory conditions (20 °C) on nematode growth medium (NGM) agar containing OP50 *Escherichia coli*. For experiments involving *glp-1(e2141)*, wild-type and mutant strains were grown during development at 25 °C. For ultraviolet treatment, worms were irradiated with 310–315 nm UVB light using narrow-band TL01 36-W bulbs in a Waldmann UV 181 BL irradiation device or were left untreated. For ionizing radiation treatment, worms were exposed to X-ray irradiation using ISOVOLT, Titan E machine from GE with a 0.5-mm aluminium filter. For heat stress treatment worms were placed on 35-mm dishes containing NGM agar and bacteria, packed into closed carton boxes to avoid dehydration and put into the Sanyo MIR-154 incubator that was running at 35 °C.

**Full Methods** and any associated references are available in the online version of the paper.

Received 4 October 2012; accepted 12 July 2013.

Published online 25 August 2013.

- Bartek, J. & Lukas, J. DNA damage checkpoints: from initiation to recovery or adaptation. *Curr. Opin. Cell Biol.* **19**, 238–245 (2007).
- Gartner, A., Milstein, S., Ahmed, S., Hodgkin, J. & Hengartner, M. O. A conserved checkpoint pathway mediates DNA damage–induced apoptosis and cell cycle arrest in *C. elegans*. *Mol. Cell* **5**, 435–443 (2000).
- Johnson, T. E. & Hartman, P. S. Radiation effects on life span in *Caenorhabditis elegans*. *J. Gerontol.* **43**, B137–B141 (1988).
- Lans, H. *et al.* Involvement of global genome repair, transcription coupled repair, and chromatin remodeling in UV DNA damage response changes during development. *PLoS Genet.* **6**, e1000941 (2010).
- Hsu, A. L., Murphy, C. T. & Kenyon, C. Regulation of aging and age-related disease by DAF-16 and heat-shock factor. *Science* **300**, 1142–1145 (2003).
- Garcia-Muse, T. & Boulton, S. J. Distinct modes of ATR activation after replication stress and DNA double-strand breaks in *Caenorhabditis elegans*. *EMBO J.* **24**, 4345–4355 (2005).

- Moser, S. C. *et al.* Functional dissection of *Caenorhabditis elegans* CLK-2/TEL2 cell cycle defects during embryogenesis and germline development. *PLoS Genet.* **5**, e1000451 (2009).
- Colaiácovo, M. P. *et al.* Synaptonemal complex assembly in *C. elegans* is dispensable for loading strand-exchange proteins but critical for proper completion of recombination. *Dev. Cell* **5**, 463–474 (2003).
- Lin, K., Hsin, H., Libina, N. & Kenyon, C. Regulation of the *Caenorhabditis elegans* longevity protein DAF-16 by insulin/IGF-1 and germline signaling. *Nature Genet.* **28**, 139–145 (2001).
- Yamawaki, T. M. *et al.* The somatic reproductive tissues of *C. elegans* promote longevity through steroid hormone signaling. *PLoS Biol.* **8**, e1000468 (2010).
- Rutkowski, R. *et al.* Regulation of *Caenorhabditis elegans* p53/CEP-1-dependent germ cell apoptosis by Ras/MAPK signaling. *PLoS Genet.* **7**, e1002238 (2011).
- Karpac, J., Younger, A. & Jasper, H. Dynamic coordination of innate immune signaling and insulin signaling regulates systemic responses to localized DNA damage. *Dev. Cell* **20**, 841–854 (2011).
- Dent, P., Yacoub, A., Fisher, P. B., Hagan, M. P. & Grant, S. MAPK pathways in radiation responses. *Oncogene* **22**, 5885–5896 (2003).
- Greiss, S., Schumacher, B., Grandien, K., Rothblatt, J. & Gartner, A. Transcriptional profiling in *C. elegans* suggests DNA damage dependent apoptosis as an ancient function of the p53 family. *BMC Genomics* **9**, 334 (2008).
- Shivers, R. P., Youngman, M. J. & Kim, D. H. Transcriptional responses to pathogens in *Caenorhabditis elegans*. *Curr. Opin. Microbiol.* **11**, 251–256 (2008).
- Nicholas, H. R. & Hodgkin, J. The ERK MAP kinase cascade mediates tail swelling and a protective response to rectal infection in *C. elegans*. *Curr. Biol.* **14**, 1256–1261 (2004).
- O'Rourke, D., Baban, D., Demidova, M., Mott, R. & Hodgkin, J. Genomic clusters, putative pathogen recognition molecules, and antimicrobial genes are induced by infection of *C. elegans* with *M. nematophilum*. *Genome Res.* **16**, 1005–1016 (2006).
- Troemel, E. R. *et al.* p38 MAPK regulates expression of immune response genes and contributes to longevity in *C. elegans*. *PLoS Genet.* **2**, e183 (2006).
- Irazoqui, J. E., Urbach, J. M. & Ausubel, F. M. Evolution of host innate defence: insights from *Caenorhabditis elegans* and primitive invertebrates. *Nature Rev. Immunol.* **10**, 47–58 (2010).
- Thomas, J. H. Concerted evolution of two novel protein families in *Caenorhabditis* species. *Genetics* **172**, 2269–2281 (2006).
- Shivers, R. P. *et al.* Phosphorylation of the conserved transcription factor ATF-7 by PMK-1 p38 MAPK regulates innate immunity in *Caenorhabditis elegans*. *PLoS Genet.* **6**, e1000892 (2010).
- Shivers, R. P., Kooistra, T., Chu, S. W., Pagano, D. J. & Kim, D. H. Tissue-specific activities of an immune signaling module regulate physiological responses to pathogenic and nutritional bacteria in *C. elegans*. *Cell Host Microbe* **6**, 321–330 (2009).
- Segref, A., Torres, S. & Hoppe, T. A screenable *in vivo* assay to study proteostasis networks in *Caenorhabditis elegans*. *Genetics* **187**, 1235–1240 (2011).



24. Muruve, D. A. *et al.* The inflammasome recognizes cytosolic microbial and host DNA and triggers an innate immune response. *Nature* **452**, 103–107 (2008).
25. Gasser, S., Orsulic, S., Brown, E. J. & Raulet, D. H. The DNA damage pathway regulates innate immune system ligands of the NKG2D receptor. *Nature* **436**, 1186–1190 (2005).
26. Rodier, F. *et al.* Persistent DNA damage signalling triggers senescence-associated inflammatory cytokine secretion. *Nature Cell Biol.* **11**, 973–979 (2009); erratum **11**, 973–979 (2009).
27. Iizuka, M. & Konno, S. Wound healing of intestinal epithelial cells. *World J. Gastroenterol.* **11**, 1272 (2009).
28. Gregorio, J. *et al.* Plasmacytoid dendritic cells sense skin injury and promote wound healing through type I interferons. *J. Exp. Med.* **207**, 2921–2930 (2010).
29. Richardson, C. E., Kooistra, T. & Kim, D. H. An essential role for XBP-1 in host protection against immune activation in *C. elegans*. *Nature* **463**, 1092–1095 (2010).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank S. Torres for technical support, P. Frommolt for advice on statistics, and A. Williams for comments on the manuscript. *C. elegans* strains were kindly provided by the CGC (funded by the US National Institutes of Health (NIH) Office of Research Infrastructure Programs (P40 OD010440)), and the Mitani laboratory. We

thank D. Kim, F. Ausubel and V. Jantsch for strains and reagents. M.A.E. received the European Molecular Biology Organization (EMBO) long-term fellowship, A.D. received the IGS-DHD fellowship, and H.-L.O. received the CECAD fellowship. O.U. acknowledges funding from the Deutsche Forschungsgemeinschaft (DFG) (SFB 670-TP4), and T.H. from the EC Network of Excellence RUBICON (LSHC-CT-2005-018683), DFG (CECAD, FOR885, SFB635, KFO286 and HO2541/4-1). B.S. acknowledges funding from the DFG (CECAD, SFB 829 and KFO 286), European Research Council (starting grant 260383), Marie Curie (FP7 ITN CodeAge 316354, aDDress 316390, MARRIAGE 316964 and ERG 239330), German–Israeli Foundation (GIF 2213-1935.13/2008 and 1104-68.11/2010), Deutsche Krebshilfe (109453) and BMBF (SyBaCol).

**Author Contributions** M.A.E. and B.S. designed the study and analysed data; M.A.E., A.S., A.D., H.-L.O. and J.I.S. performed experiments and discussed data; A.S. and T.H. designed UPS-specific experiments; O.U. provided advice and reagents; and B.S. wrote the manuscript.

**Author Information** Data are deposited at ArrayExpress under the accession number E-MTAB-1689. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to B.S. ([bjoern.schumacher@uni-koeln.de](mailto:bjoern.schumacher@uni-koeln.de)).

## METHODS

**Worm strains.** All strains were cultured according to standard conditions<sup>30</sup>. Strains used were N2 (Bristol, wild type), FX03886 *xpc-1(tm3886)*, DR26 *daf-16(m26)*, AA67 *daf-12(rh61 rh411);daf-16(mgDf50)*, DW101 *atl-1(tm853)*, SP506 *rad-5(mn159)*, PS3551 *hsf-1(sy441)*, *atl-1(tm853);atm-1(gk186)*, WS2277 *hus-1(op241)*, XY1054 *cep-1(lg12501)*, MT1522 *ced-3(n717)*, MT8186 *mpk-1(oz140)*, VC8 *jnk-1(gk7)*, KU25 *pmk-1(km25)*, CB4037 *glp-1(e2141)*, AV157 *spo-11(me44)/nT1(IV;V)*, AV276 *syp-2(ok307)/nT1 [qIs51] (V)*, *spo-11(ok79) IV/nT1;syp-2(ok307) V/nT1*, AU78 *agIs219* (carrying PT24B8.5::GFP::unc-54-3'UTR), ZD442 *agIs219 atf-7(qd22) III*, ZD318 *agIs219 atf-7(qd22 qd130) III*, BC10060 *dpy-5(e907);sEx884[hsp-70::GFP + pCeh361]*, PP563 *unc-119(ed4);hhIs64[unc-119(+);sur-5::UbV-GFP]III*, PP545 *unc-119(ed4);hhIs53[unc-119(+);sur-5::K29/48R-UbV-GFP]*, PP556 *unc-119(ed4);hhIs57[unc-119(+);sur-5::GFP]*.

**Bacterial strains and growth conditions.** *E. coli* (OP50) and *Pseudomonas aeruginosa* (UCBPP-PA14) were seeded on lysogeny broth agar and kept at 4 °C. *Bacillus subtilis* (ATCC9372) was grown on blood agar plates and kept at the same temperature. Overnight cultures were grown at 37 °C in lysogeny broth media without antibiotics in all cases. To feed worms with *P. aeruginosa*, enriched peptone plates were used as described previously<sup>31</sup>; in all other cases feeding was performed on standard matode growth medium (NGM) agar.

**Microarray analysis.** Approximately 1,000 age-synchronized young adult hermaphrodites (24 h post the L4 larval stage) per sample were irradiated with 0 mJ cm<sup>-2</sup> or 60 mJ cm<sup>-2</sup> of ultraviolet (three different sample per condition). RNA was extracted 6 h post treatment using TRIzol (Invitrogen) according to the manufacturer's protocol. RNA was purified from TRIzol using the RNeasy Mini Kit (Qiagen). RNA quality was assessed using a Nanodrop 1000 spectrophotometer (Thermo Scientific) and a Bioanalyzer (Agilent). Expression profiles were obtained using a GeneChip *C. elegans* Genome Array (Affymetrix) according to the manufacturer's specifications at the Cologne Centre for Genomics (CCG) facility. Hybridization signals were normalized by quantile normalization and differentially expressed genes were determined by two-way analysis of variance (ANOVA) using the Partek software package.

**Gene Ontology classification and secreted peptide analysis.** All entries of significantly differentially expressed genes were subjected to Gene Ontology classification (<http://www.geneontology.org>). Significant overrepresentation of Gene-Ontology-classified biological processes was assessed by comparing the number of pertinent genes in a given biological process to the total number of the relevant genes printed on the array for that particular biological process (Fisher exact test) using the publicly accessible software DAVID (<http://david.abcc.ncifcrf.gov/summary.jsp>). Owing to the redundant nature of Gene Ontology annotations, we used Kappa statistics to measure the degree of the common genes between two annotations, and heuristic clustering to classify the groups of similar annotations according to kappa values.

Putatively secreted peptides were analysed for the presence of signal peptides using SignalP (<http://www.cbs.dtu.dk/services/SignalP>) as described in ref. 32 and for target compartments using WoLF PSORT (<http://wolfsort.org/>) as described in ref. 33.

**Real-time quantitative PCR analysis.** For qPCR analysis, complementary DNA was generated using Superscript II (Invitrogen). Quantitative real-time PCR (qPCR) was carried out using Biorad MyIQ real-time PCR machines using SYBR Green I (Sigma) and Platinum Taq polymerase (Invitrogen). All qPCR reactions were carried out in duplicate. The generation of specific PCR products was confirmed by melting curve analysis. Each primer pair was tested with a logarithmic dilution of a cDNA mix to generate a linear standard curve, crossing point (CP) plotted versus log of template concentration, which was used to calculate the primer pair efficiency ( $E = 10^{(-1/\text{slope})}$ ). For data analysis, the second derivative maximum method was applied, and induction of target cDNA was calculated using the equation:

$$(E_{\text{target}})^{\Delta \text{CP}(\text{cDNA}_{\text{untreated}} - \text{cDNA}_{\text{treated}})_{\text{target}}} / (E_{\text{control}})^{\Delta \text{CP}(\text{cDNA}_{\text{untreated}} - \text{cDNA}_{\text{treated}})_{\text{control}}}$$

For GFP expression, worms were analysed as described previously<sup>34</sup>, and for the expression of immune genes the analysis was performed as described in ref. 21.

Expression levels were normalized to actin (*act-1*), lamin (*lmn-1*), and g-Tubulin (*tbg-1*).

**Irradiation.** For ultraviolet treatment, worms were irradiated with 310–315-nm UVB light using narrow band TL01 36-W bulbs in a Waldmann UV 181 BL irradiation device or were left untreated. Irradiance was measured using a UVX digital radiometer and a UVX-31 probe from UVP. For ionizing radiation treatment, worms were exposed to X-ray irradiation using ISOVOLT, Titan E machine from GE with a 0.5-mm Aluminium filter.

**Fecundity.** Worms were staged and L4 larvae were irradiated and allowed to recover for 24 h at 20 °C. Worms were transferred to new OP50-seeded plates and allowed to lay eggs. Adult worms were transferred to new OP50-seeded plates every 24 h until the end of the experiment. Viable offspring on each plate was counted 48 h later.

**Heat stress treatment.** Worms were placed on 35-mm dishes containing NGM agar and bacteria at a density of 35 worms per plate. For each condition, three independent plates were used. Plates were packed into closed carton boxes to avoid dehydration and put into the Sanyo MIR-154 incubator that was running at 35 °C. Experiments involving *P. aeruginosa* were performed on 65-mm dishes; two dishes per condition, with 55 worms per dish. This was done owing to technical limitations of the S2 facility where pathogenic materials were handled.

Worms with protruding vulva were excluded from the survival experiments, which is a standard procedure in *C. elegans* survival studies.

**Statistical analysis.** Error bars represent standard deviations in all cases. Statistical significance was determined by using log rank analysis for all stress resistance and pathogen resistance experiments. The survival curves for the treatment conditions do not cross each other at any time point. For progeny survival, paired *t*-test with two-tailed distribution was performed to evaluate statistical significance. Pearson correlation analysis was performed using the Prism software package.

For most experiments *n* was over 100. Randomization was not applied because the group allocation was guided by the genotype of the respective mutant worms. Worms of a given genotype were randomly selected from large strain populations for each experiment without any preconditioning. Blinding was not applied as the experiments were carried out under highly standardized and predefined conditions such that an investigator-induced bias can be excluded.

**Microscopy.** GFP-expressing worms were analysed and images were taken by using a LEICA M165FC microscope and LEICA Application Suite V3.3.0 software.

**Western blot analysis.** Immunoblot analysis was performed according to a protocol described previously<sup>23</sup>. Antibodies used were anti-GFP Living Colours A.v. Monoclonal Antibody (JL-8), Clontech art-no 632381; Mouse Monoclonal anti- $\alpha$ Tubulin (Clone DM 1A), Sigma-Aldrich art-no T6199; Phospho-p44/42 MAPK (Erk1/2) (Thr202/Tyr204) Antibody 9101, Cell Signalling, NEB art-no 9101S.

**Western blot quantification.** Western blot quantification was carried out using fluorescently labelled IRDye secondary antibodies, an Odyssey infrared scanner and Odyssey V3.0 software.

**RNA interference treatment.** HT115 bacteria containing specific RNAi constructs were grown on lysogeny broth agar plates supplemented with ampicillin and tetracycline. Plates were kept at 4 °C. Overnight cultures were grown in lysogeny broth media containing ampicillin. RNAi expression was induced by adding 1 mM isopropylthiogalactoside (IPTG) and incubating the cultures at 37 °C for 20 min before seeding bacteria on NGM agar supplemented with ampicillin and 3 mM IPTG.

30. Brenner, S. The genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71–94 (1974).
31. Kim, D. H. *et al.* A conserved p38 MAP kinase pathway in *Caenorhabditis elegans* innate immunity. *Science* **297**, 623–626 (2002).
32. Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods* **8**, 785–786 (2011).
33. Horton, P. *et al.* WoLF PSORT: protein localization predictor. *Nucleic Acids Res.* **35**, W585–W587 (2007).
34. Bush, K. T., Goldberg, A. L. & Nigam, S. K. Proteasome inhibition leads to a heat-shock response, induction of endoplasmic reticulum chaperones, and thermotolerance. *J. Biol. Chem.* **272**, 9086–9092 (1997).

# Synthetic lethal metabolic targeting of cellular senescence in cancer therapy

Jan R. Dörr<sup>1</sup>, Yong Yu<sup>2</sup>, Maja Milanovic<sup>1</sup>, Gregor Beuster<sup>2</sup>, Christin Zasada<sup>3</sup>, J. Henry M. Däbritz<sup>1</sup>, Jan Lisec<sup>1,4</sup>, Dido Lenze<sup>5</sup>, Anne Gerhardt<sup>1</sup>, Katharina Schleicher<sup>1</sup>, Susanne Kratzat<sup>6</sup>, Bettina Purfürst<sup>2</sup>, Stefan Walenta<sup>7</sup>, Wolfgang Mueller-Klieser<sup>7</sup>, Markus Gräler<sup>8</sup>, Michael Hummel<sup>5</sup>, Ulrich Keller<sup>6</sup>, Andreas K. Buck<sup>9</sup>, Bernd Dörken<sup>1,2</sup>, Lothar Willmitzer<sup>10</sup>, Maurice Reimann<sup>1</sup>, Stefan Kempa<sup>3</sup>, Soyoung Lee<sup>1,2</sup> & Clemens A. Schmitt<sup>1,2</sup>

**Activated oncogenes and anticancer chemotherapy induce cellular senescence, a terminal growth arrest of viable cells characterized by S-phase entry-blocking histone 3 lysine 9 trimethylation (H3K9me3)<sup>1,2</sup>. Although therapy-induced senescence (TIS) improves long-term outcomes<sup>3</sup>, potentially harmful properties of senescent tumour cells make their quantitative elimination a therapeutic priority. Here we use the Eμ-myc transgenic mouse lymphoma model in which TIS depends on the H3K9 histone methyltransferase Suv39h1 to show the mechanism and therapeutic exploitation of senescence-related metabolic reprogramming *in vitro* and *in vivo*. After senescence-inducing chemotherapy, TIS-competent lymphomas but not TIS-incompetent *Suv39h1*<sup>−</sup> lymphomas show increased glucose utilization and much higher ATP production. We demonstrate that this is linked to massive proteotoxic stress, which is a consequence of the senescence-associated secretory phenotype (SASP) described previously<sup>4–6</sup>. SASP-producing TIS cells exhibited endoplasmic reticulum stress, an unfolded protein response (UPR), and increased ubiquitination, thereby targeting toxic proteins for autophagy in an acutely energy-consuming fashion. Accordingly, TIS lymphomas, unlike senescence models that lack a strong SASP response, were more sensitive to blocking glucose utilization or autophagy, which led to their selective elimination through caspase-12- and caspase-3-mediated endoplasmic-reticulum-related apoptosis. Consequently, pharmacological targeting of these metabolic demands on TIS induction *in vivo* prompted tumour regression and improved treatment outcomes further. These findings unveil the hypercatabolic nature of TIS that is therapeutically exploitable by synthetic lethal metabolic targeting.**

Anticancer agents are known to induce cellular senescence, but diagnostic procedures to non-invasively detect a senescent proliferation arrest of cancer lesions in response to chemotherapy *in vivo* remain to be established<sup>7–10</sup>. As a well-characterized model system, we used primary Eμ-myc lymphomas that were stably transduced with the anti-apoptotic *Bcl2* moiety to investigate TIS in the absence of an apoptotic response<sup>3</sup>. Lymphoma cells were transplanted into recipient mice, and treated when enlarged lymph nodes became palpable. A 5-day exposure of tumour-bearing mice to the DNA-damaging anticancer drug cyclophosphamide (CTX) rendered *Suv39h1*-proficient lymphomas (hereafter referred to as 'control'), but not *Suv39h1*-deficient lymphomas, uniformly senescent (Fig. 1a). In contrast to control;*Bcl2* lymphomas, *Suv39h1*<sup>−</sup>;*Bcl2* lymphomas not only remained negative in terms of senescence-associated β-galactosidase (SA-β-gal) activity<sup>11</sup> but also preserved S-phase activity, as indicated by the retained incorporation of the thymidine analogue 5-bromodeoxyuridine (BrdU) *in situ*. *In vitro* exposure to adriamycin (ADR), another chemotherapeutic agent,

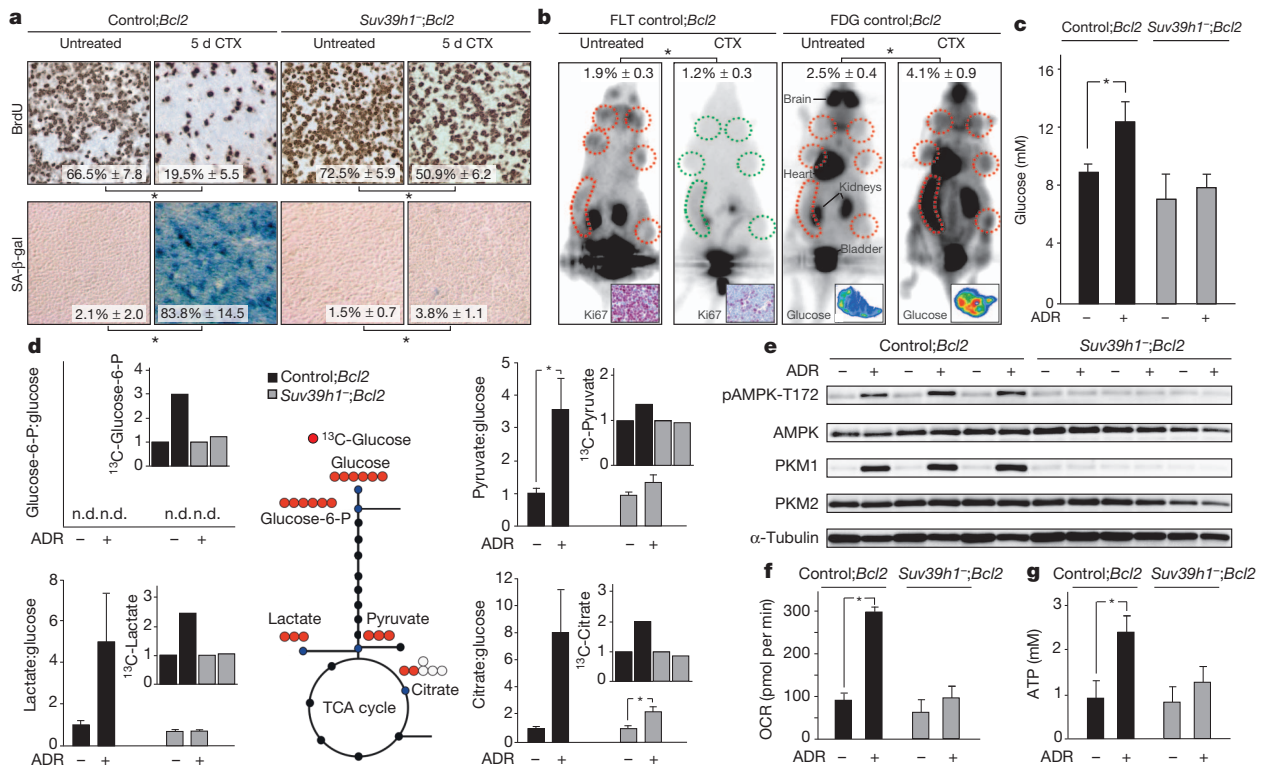
also produced *Suv39h1*-dependent senescence, detectable even in lymphoma cells without *Bcl2* overexpression (Supplementary Fig. 1). Analogous to BrdU incorporation *in vitro*, we applied [<sup>18</sup>F] fluorothymidine (FLT) positron emission tomography (PET) to visualize cell growth *in vivo*. Consistent with TIS as assessed by SA-β-gal staining and the proliferation markers BrdU and Ki67 *in situ*, serial imaging analyses of lymphoma-bearing mice before and 6 days after CTX by FLT-PET demonstrated significantly lower FLT tracer uptake after chemotherapy in the control;*Bcl2* group only (Fig. 1b, and Supplementary Fig. 2a). Unexpectedly, the same drug-senescent control;*Bcl2* lymphomas showed an enhanced glucose uptake by a [<sup>18</sup>F] fluorodeoxyglucose (FDG)-PET, and presented with higher glucose levels *in situ* compared to the untreated, proliferating condition, whereas no difference was seen in TIS-incompetent *Suv39h1*<sup>−</sup>;*Bcl2* lymphomas before and after therapy. Hence, TIS lymphomas can be non-invasively detected and are characterized by a discordant negative FLT-PET but positive FDG-PET scan post treatment. Notably, a positive FDG-PET often predicts an imminent clinical progress or relapse, however, lastingly stabilized diseases are not uncommon in patients presenting with FDG-PET-positive lesions after chemotherapy<sup>12</sup>.

To elucidate the underlying metabolic alterations in TIS cells, we first confirmed the senescence-related increased glucose avidity in a PET-independent 2-(N-(7-nitrobenz-2-oxa-1,3-diazol-4-yl)amino)-2-deoxyglucose (2NBDG) uptake assay after *in vitro* exposure to ADR (Supplementary Fig. 2b). Moreover, only senescence-capable cells exhibited increased glucose transporter and glycolytic enzyme expression levels after treatment (Supplementary Fig. 2c). As a result, steady-state glucose concentrations were much higher in ADR-senescent cells as compared to untreated, proliferating control;*Bcl2* lymphoma cells, or to *Suv39h1*<sup>−</sup>;*Bcl2* lymphoma cells, irrespective of treatment (Fig. 1c and Supplementary Fig. 2d). TIS cells metabolized glucose rapidly to pyruvate, lactate and citrate, such that these metabolites became detectable at much higher concentrations relative to glucose, despite the resting state of these cells (Fig. 1d and Supplementary Fig. 2e, f). Only ADR-senescent cells showed phosphorylation activation of the adenosine mono-phosphate (AMP)-activated protein kinase (AMPK), a central energy sensor of the cell, indicating their increased energetic demands<sup>13</sup> (Fig. 1e and Supplementary Fig. 2g). These results are consistent with a TIS-enforced enhancement of the 'Warburg effect', the predominantly non-oxidative breakdown of glucose typically detectable in tumour cells<sup>14</sup>.

The increased glycolytic activity in TIS lymphomas was associated with induced levels of the M1 isoform of pyruvate kinase (PKM1), which promotes further utilization of pyruvate in the tri-carboxylic

<sup>1</sup>Charité-Universitätsmedizin Berlin, Molekulares Krebsforschungszentrum (MKFZ), Augustenburger Platz 1, 13353 Berlin, Germany. <sup>2</sup>Max-Delbrück-Center for Molecular Medicine (MDC), Robert-Rössle-Straße 10, 13125 Berlin, Germany. <sup>3</sup>Integrative Metabolomics and Proteomics, Berlin Institute of Medical Systems Biology, Max-Delbrück-Center for Molecular Medicine, Robert-Rössle-Straße 10, 13125 Berlin, Germany. <sup>4</sup>German Cancer Consortium, Deutsches Krebsforschungszentrum (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany. <sup>5</sup>Charité-Universitätsmedizin Berlin, Department of Pathology, Charitéplatz 1, Berlin 10117, Germany. <sup>6</sup>Technische Universität München, III Medical Department, Ismaninger Straße 22, 81675 Munich, Germany. <sup>7</sup>Universitätsmedizin der Johannes Gutenberg-Universität, Institute of Physiology and Pathophysiology, Duesbergweg 6, 55128 Mainz, Germany. <sup>8</sup>Universitätsklinikum Jena, Department of Anesthesiology and Intensive Care Medicine & Center for Sepsis Control and Care (CSCC), Erlanger Allee 1, 07747 Jena, Germany. <sup>9</sup>Universitätsklinikum Würzburg, Department of Nuclear Medicine, Oberdürrbacher Straße 6, 97080 Würzburg, Germany. <sup>10</sup>Max Planck Institute of Molecular Plant Physiology, Am Mühlenberg 1, 14476 Potsdam, Germany.





**Figure 1 | TIS is a highly glucose-consuming and energy-producing condition.** **a**, BrdU incorporation reflecting DNA synthesis and SA-β-gal activity as markers of CTX-induced senescence in Eμ-*myc* transgenic control;*Bcl2* versus *Suv39h1*<sup>-/-</sup>;*Bcl2* lymphomas *in vivo*. Numbers represent mean percentages of positive cells ± s.d. ( $n = 5$  each). **b**, Representative whole-body FLT- and FDG-PET scans of an individual control;*Bcl2* lymphoma-bearing mouse at tumour manifestation (untreated) and at day 6 after a single dose of CTX. Red circles, areas of specific tracer retention; green circles, matching areas with lost retentions. Specific tracer uptake was assessed by calculation of tumour-to-background ratios. Data represent mean ratios ± s.d. ( $n = 5$  mice for FLT-PET scans and  $n = 6$  mice for FDG-PET scans). Insets show representative Ki67 staining and colour-calibrated (red, highest intensities; blue, lowest intensities) bioluminescence images of glucose levels in tissue sections of the respective groups. **c**, Intracellular glucose concentrations in untreated and ADR-treated control;*Bcl2* (that is, senescent) versus *Suv39h1*<sup>-/-</sup>;*Bcl2* (that is, non-senescent) lymphomas *in vitro* (as in **d–g**). Results represent means ± s.d. ( $n = 5$  each). **d**, Extracellular levels of the specified metabolite-to-glucose ratios in the indicated lymphoma cells groups relative to

acid (TCA) cycle<sup>15</sup> (Fig. 1d, e). Indeed, we detected a strongly elevated oxygen consumption rate (OCR) and much higher intracellular ATP levels selectively in ADR-senescent lymphoma cells (Fig. 1f, g and Supplementary Fig. 2h; see also Supplementary Fig. 3 for similar findings in murine and human fibroblasts undergoing oncogene-induced senescence (OIS)). Lymphoma metabolomics further underscored the global catabolic nature of senescence-associated metabolic alterations (Supplementary Fig. 4). Taken together, drug-senescent lymphomas are characterized by an unexpected hypermetabolic phenotype composed of enhanced glycolysis, fatty acid catabolism and ATP-boosting oxidative phosphorylation. Notably, enhanced energy production was found as the common denominator of all TIS lymphomas tested, which may differ in their individual utilization of energy-generating metabolic pathways (Supplementary Fig. 5).

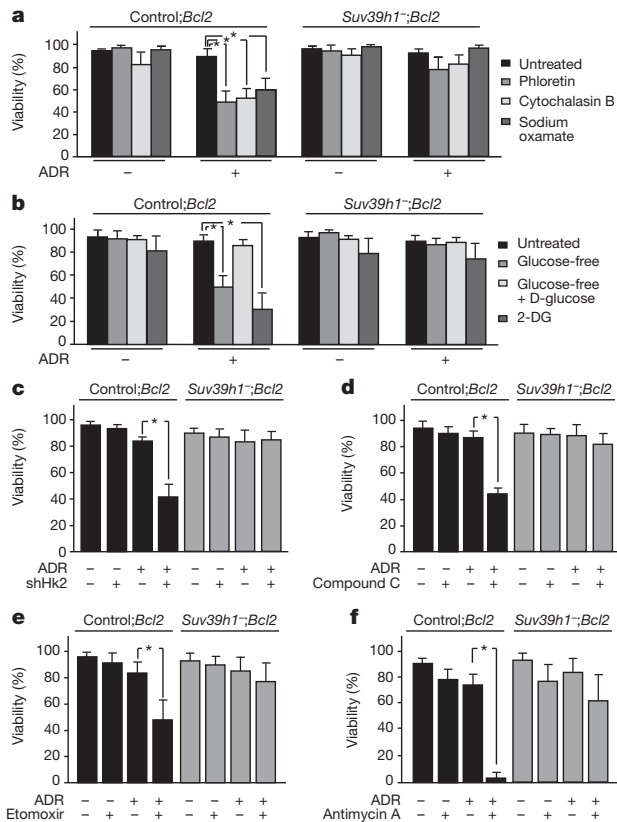
We next asked whether TIS lymphomas not only show a hypermetabolic condition but also depend on it. Indeed, senescent cells were selectively susceptible to inhibition of glucose transporters (that is, phloretin or cytochalasin B), to a glycolytic block by genetic (that is, stable knockdown of hexokinase 2) or pharmacological (that is, 2-deoxy-D-glucose (2DG)

untreated control;*Bcl2* lymphoma cells after 18 h of cultivation. Data represent mean ratios ± s.d. ( $n = 4$  each, measurements were carried out in duplicate). Note that glucose-6-phosphate is not detectable (n.d.) in the medium. Insets show (ADR-treated relative to untreated) intracellular levels of the respective metabolites after labelling with <sup>13</sup>C<sub>6</sub>-glucose for one representative lymphoma from each group. Data represent mean ratios (see Supplementary Fig. 5 for an assessment of the heterogeneity of the TIS-associated hypermetabolism and a comparison of the metabolic parameters of individual lymphomas).

**e**, Immunoblot analyses of phosphorylated AMP-activated protein kinase (pAMPK), total AMPK, pyruvate kinase isoforms M1 and M2 in ADR-senescent control;*Bcl2* versus ADR-treated *Suv39h1*<sup>-/-</sup>;*Bcl2* lymphoma cells *in vitro* ( $n = 3$  each; α-tubulin as a loading control). Gels were cropped to ease presentation; full-length blots are provided in Supplementary Fig. 16). **f**, Mean OCRs in ADR-senescent control;*Bcl2* versus ADR-treated *Suv39h1*<sup>-/-</sup>;*Bcl2* lymphoma cells *in vitro*. Results represent means ± s.d. ( $n = 5$  each). **g**, Intracellular ATP concentrations in cells as in **f**. Results represent means ± s.d. ( $n = 5$  each). \* $P < 0.05$ .

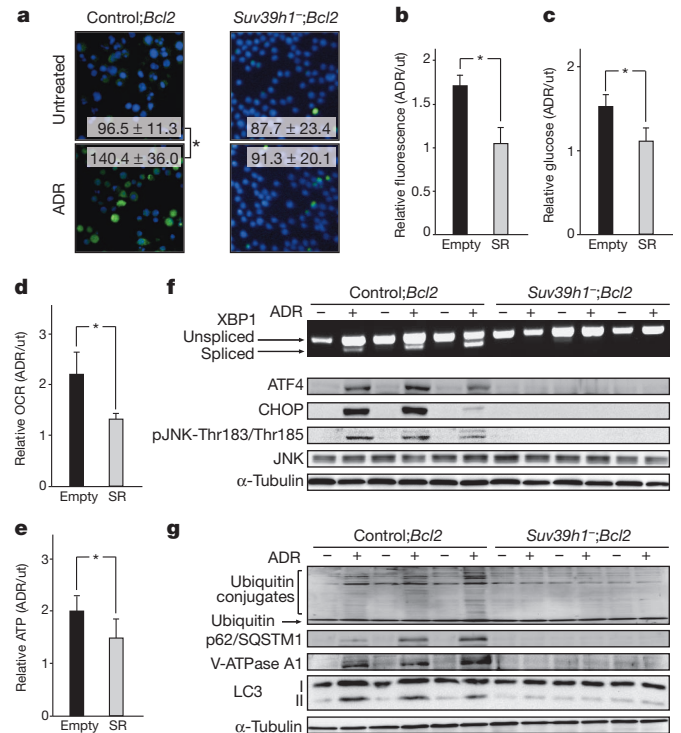
as a false substrate, and sodium oxamate as an inhibitor of the lactate dehydrogenase) means, or to compound C-mediated inhibition of the energy sensor AMPK (Fig. 2a–d and Supplementary Fig. 6). TIS-selective toxicity was also observed after suppressing fatty acid oxidation by etomoxir, or ATP depletion after poisoning of the respiratory chain by antimycin A (Fig. 2e, f). Thus, TIS tumour cells rely on energy-providing hypercatabolism, thereby establishing a critical senescence-associated liability that may be therapeutically exploitable.

Recently, enhanced uptake of amino acids and higher protein quantities in Ras-senescent fibroblasts have been reported<sup>16</sup>, which may be explained, at least to some extent, by SASP, the markedly induced synthesis of largely nuclear factor-κB (NF-κB)-dependent pro-inflammatory cytokines in OIS and TIS<sup>4–6,17,18</sup>. Consistent with this view, we also found a significantly higher protein synthesis rate in TIS lymphoma cells (Fig. 3a). Accordingly, gene set enrichment analyses (GSEA) were strongly skewed towards the TIS group for transcripts linked to protein maturation and protein processing (Supplementary Fig. 7a). Further supporting a link between hypermetabolism and senescence-associated cytokine production, we found senescent NF-κB-inactivated Eμ-*myc*;*Bcl2* lymphomas



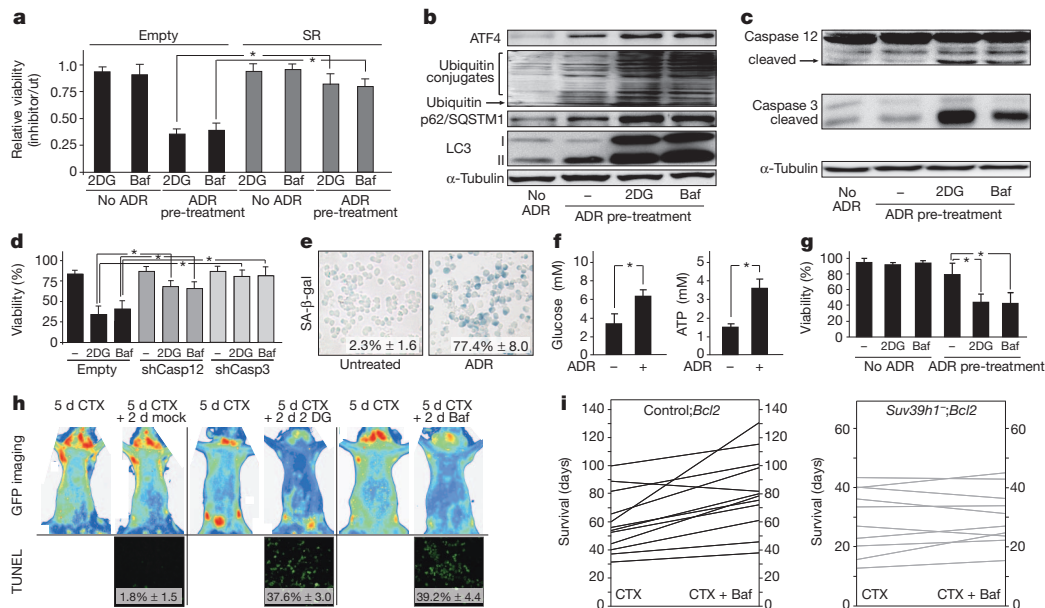
**Figure 2 | Selective vulnerability of TIS cells to inhibition of energy-generating catabolic pathways.** **a–e**, Cell viability by trypan blue dye exclusion in ADR-senescent control;*Bcl2* versus ADR-treated *Suv39h1<sup>-/-</sup>;Bcl2* lymphoma cells at day 5 compared to the untreated condition in various biochemical, pharmacological or genetic interventional settings *in vitro*. All results in this figure represent means  $\pm$  s.d. ( $n = 5$  each). **a**, Inhibition of glucose metabolism by phloretin, cytochalasin B, and sodium oxamate. **b**, Exposure to 2DG or switch to glucose-free medium with or without glucose replenishment. **c**, Lentiviral knockdown of hexokinase 2 (Hk2) by short hairpin RNA (shHk2). **d**, Inhibition of AMPK by compound C. **e**, Block of fatty acid oxidation by the carnitine palmitoyltransferase I inhibitor etomoxir. **f**, Inhibition of the electron transport chain by antimycin A. \* $P < 0.01$ , except **b** ( $n = 8$  each) and **e**, in which \* $P < 0.05$ .

(stably expressing the NF- $\kappa$ B super-repressor I $\kappa$ B $\Delta$ N; Supplementary Fig. 7b), which indeed produce protein only slightly above the baseline of non-senescent NF- $\kappa$ B-intact cells (and specifically express much less SASP cytokines<sup>18</sup>), and showed much lower increases in glucose uptake, OCR and ATP generation (Fig. 3b–e). As the underlying mechanism, we proposed that the high and presumably not always proper production and processing of senescence-associated secretable peptides may impose substantial proteotoxic stress, in turn evoking energy-consuming countermeasures for cell survival. In fact, we detected ultrastructural endoplasmic reticulum alterations reminiscent of endoplasmic reticulum stress, a significant enrichment of transcripts involved in the UPR, as well as UPR-related phosphorylation of Jun N-terminal kinases (JNK), alternative splicing of *XBP1* transcripts, and enhanced expression of the transcription factors ATF4 (cyclic AMP-dependent transcription factor ATF-4) and CHOP (also known as DNA damage-inducible transcript 3 protein) selectively in TIS lymphomas, all of which indicate proteotoxic stress<sup>19</sup> (Fig. 3f and Supplementary Fig. 7c). Moreover, consistent with the ubiquitin-mediated degradation of endoplasmic-reticulum-associated misfolded proteins through autophagy<sup>20</sup>, global ubiquitination and expression of the microtubule-associated protein light chain 3 (MAP1-LC3)-binding protein SQSTM1 (also known as p62) were induced strongly in TIS lymphomas, but to a much lesser extent in NF- $\kappa$ B-inactivated senescent lymphomas, or



**Figure 3 | TIS shows a proteotoxic stress cascade.** **a–g**, Proteostasis analysis in TIS control versus senescence-incapable lymphoma cells *in vitro*, as in Fig. 2, or versus SASP-attenuated TIS lymphoma cells. **a**, Global protein synthesis rate, measured by incorporation of green-fluorescent methionine. Mean fluorescent intensities reflect newly synthesized protein within 2 h (DAPI used as a counterstain). Data represent means  $\pm$  s.d. ( $n = 5$  each). **b**, Rate of global protein synthesis (measured as in **a**) in ADR-senescent relative to untreated (ut) control;*Bcl2* lymphoma cells; either empty vector-infected cells or cells stably expressing the NF- $\kappa$ B super-repressor I $\kappa$ B $\Delta$ N (SR, ref. 18). Results represent means  $\pm$  s.d. ( $n = 5$  each). **c**, Relative intracellular glucose concentrations in cells as in **b**. **d**, Relative OCR in cells as in **b**. **e**, Relative intracellular ATP concentrations in cells as in **b**. **f**, Expression of the transcription factors *Xbp1* (by RT-PCR (PCR with reverse transcription; top); note that the lower band reflects the UPR-related splice form), ATF4, CHOP, phospho-Jun N-terminal kinase (JNK-P-Thr183/Thr185) and total JNK (by immunoblot analysis, with  $\alpha$ -tubulin as a loading control; bottom) in lymphoma cells as indicated ( $n = 3$  each; gels were cropped to ease presentation; full-length blots are provided in Supplementary Fig. 16). **g**, Global protein ubiquitination, expression of p62/SQSTM1 (whose transcript levels were also selectively induced in TIS cells by approximately 4.8-fold (data not shown)) and V-ATPase subunit A1, and conversion of MAP1-LC3-I to MAP1-LC3-II by immunoblot analysis, demonstrating autophagosome formation in lymphoma cells as in **f** ( $\alpha$ -tubulin was used as a loading control; gels were cropped to ease presentation; full-length blots are provided in Supplementary Fig. 16). \* $P < 0.05$ .

remained at virtually undetectable levels in ADR-treated *Suv39h1<sup>-/-</sup>;Bcl2* lymphomas (Fig. 3g and Supplementary Fig. 7d). We reasoned that autophagy may relieve endoplasmic reticulum stress by degrading improper, potentially toxic proteins. Indeed, TIS cells displayed a greatly expanded lysosomal compartment (creating the basis for the SA- $\beta$ -gal assay<sup>11</sup>), accompanied by structural and biochemical features of autophagy (that is, MAP1-LC3-I-isoform to MAP1-LC3-II-isoform conversion, elevated autophagic flux, and activation of the lysosomal vacuolar type H<sup>+</sup>-ATPase (V-ATPase) A1), which were virtually absent in *Suv39h1<sup>-/-</sup>;Bcl2* lymphoma cells or, likewise, in low-SASP TIS mouse lymphoma cells or low-SASP human OIS fibroblasts (Fig. 3g, Supplementary Fig. 7d–g and Supplementary Fig. 8). In essence, the acutely engaged, high-level senescence-associated secretory program seems to overwhelm the cellular capacity for accurate protein biosynthesis, post-translational processing, vesicular transport and secretion, thus leading to high proteotoxic stress and autophagy.



**Figure 4 | Synthetic lethal targeting of TIS-related metabolic liabilities.**

**a**, Cellular viability of super-repressor- or empty-vector-infected control;*Bcl2* lymphomas treated with ADR or untreated for 5 days, then exposed for 2 days to the inhibitors 2DG or bafilomycin A1, compared to no inhibitor (ut) *in vitro*. Results represent means  $\pm$  s.d. ( $n = 5$  each). **b**, Immunoblot analyses of ATF4, global protein ubiquitination, p62/SQSTM1, and MAP1-LC3-I to MAP1-LC3-II conversion in control;*Bcl2* lymphomas with no ADR treatment, in TIS cells, and in TIS cells after exposure to 2DG or bafilomycin A1 *in vitro*; representative blot with  $\alpha$ -tubulin as a loading control (gels were cropped to ease presentation; full-length blots are provided in Supplementary Fig. 16). **c**, Cleavage (activation) of caspase 12 and detection of cleaved caspase 3 by immunoblot analysis of control;*Bcl2* lymphomas as in **b**; representative blot with  $\alpha$ -tubulin as a loading control (gels were cropped to ease presentation; full-length blots are provided in Supplementary Fig. 16). **d**, Cellular viability of ADR-treated control;*Bcl2* lymphoma cells infected with short-hairpin RNAs against either caspase 12 (shCasp12) or caspase 3 (shCasp3) compared to empty vector-infected lymphoma cells treated as in **a**. Results represent means  $\pm$  s.d. ( $n = 4$  each). **e**, TIS cells in primary human acute myeloid leukaemia (AML) samples by SA- $\beta$ -gal staining of treated ( $0.1 \mu\text{g ml}^{-1}$  ADR

for 5 days) versus untreated cells *in vitro*. Data represent means  $\pm$  s.d. ( $n = 5$  each). **f**, Intracellular glucose (left) and ATP (right) concentrations in senescent AML cells as in **e**. **g**, Cellular viability of primary AML samples treated as in **e** and exposed to inhibitors as in **a**. Results represent means  $\pm$  s.d. ( $n = 5$  each). **h**, Colour-coded whole-body fluorescence imaging to visualize green fluorescent protein (GFP)-tagged tumour burden in time-courses of three representative mice harbouring CTX-senescent control;*Bcl2* lymphomas at day 5 (see Figure 1a), and two days later, after the administration of 2DG ( $n = 3$  mice), bafilomycin A1 ( $n = 4$  mice) or mock treatment ( $n = 4$  mice) starting from day 5 post CTX (red, highest intensities; blue, lowest intensities). TdT-mediated dUTP nick end labelling (assay) (TUNEL) staining of lymph node sections *in situ* to visualize apoptosis at day 7. Numbers represent means  $\pm$  s.d. ( $n = 3$ ). **i**, Overall survival of 5 day-CTX-treated control;*Bcl2* or *Suv39h1*<sup>−/−</sup>; *Bcl2* lymphoma-bearing mice randomly assigned to sequential exposure to bafilomycin A1 or mock as in **h** (control;*Bcl2*:  $n = 12$  and *Suv39h1*<sup>−/−</sup>; *Bcl2*:  $n = 10$ ), presented as matched pair analyses ( $P < 0.001$  for the control;*Bcl2* comparison;  $P = 0.221$  for the *Suv39h1*<sup>−/−</sup>; *Bcl2* comparison).  $^*P < 0.05$ .

Based on these findings, we reasoned that senescent cells rely on an intact lysosomal protein degradation machinery, which depends on an acidic intravesicular pH maintained by lysosomal ATPases, to buffer proteotoxic stress. To examine this, we exposed untreated and TIS control;*Bcl2* lymphoma cells to bafilomycin A1 or concanamycin A, specific inhibitors of lysosomal V-ATPases, or to a cocktail of lysosomal protease inhibitors. In all of these inhibitor settings, TIS-cell deaths occurred at significantly higher rates, as also observed in response to knockdown of p62/SQSTM1, indicating that proper lysosomal function is essential for the survival of senescent cells, but to a much lesser extent for that of non-senescent cells (Supplementary Fig. 9a–c; see similar findings for Ras-infected fibroblasts in Supplementary Fig. 9d). We reasoned that the substantial proteostress-related energetic demands claimed by the expanded lysosomal V-ATPase machinery and by boosted autophagic processing critically add to the energy needed to fuel the secretory program *per se* and thereby account for the observed catabolic liability of senescence. Indeed, acute V-ATPase inhibition selectively increased ATP levels in TIS lymphoma cells (Supplementary Fig. 9e). Moreover, senescence priming by oncogenic Ras resulted in a profound and Atg5-dependent increase of oxygen consumption, whereas, in contrast, SASP-attenuated rapamycin-induced senescence (which is accompanied by direct elevation of autophagic activity that does not involve energy-dissipating proteotoxicity-driven endoplasmic reticulum stress) occurred with an unchanged OCR (Supplementary Fig. 9e–k). Accordingly, senescent NF- $\kappa$ B-inactivated E $\mu$ -myc;*Bcl2* lymphomas,

like rapamycin-senescent lymphomas, were much less susceptible to 2DG- or bafilomycin-A1-evoked metabolic killing (Fig. 4a and Supplementary Fig. 9l). Importantly, inhibition of glycolysis by 2DG or of V-ATPases by bafilomycin A1 in SASP-producing TIS cells resulted in further accumulation of misfolded proteins, as indicated by additional induction of the UPR marker ATF4 and even higher protein ubiquitination, thereby presumably exceeding tolerable endoplasmic reticulum stress and triggering an endoplasmic reticulum death signal (Fig. 4b).

We next examined whether 2DG or bafilomycin A1 provokes endoplasmic-reticulum-related cell death, which is characterized by Ca<sup>2+</sup>-flux-dependent activating cleavage of caspase 12 (or its orthologue caspase 4 in human cells) followed by proteolytic processing of caspase 3 (ref. 21). Indeed, energy depletion or direct inhibition of autophagy were sufficient to activate these caspases in TIS cells and to produce caspase-12- or caspase-3-dependent cell death (Fig. 4c, d and Supplementary Fig. 10). To confirm our findings across species and tumour entities, we screened a variety of human cancer cell lines for their ability to enter TIS, and found that only TIS-capable cell lines showed increased cellular glucose and ATP levels, which correlated with selective susceptibility to 2DG and bafilomycin A1 treatment (Supplementary Fig. 11 and Supplementary Table 1). Moreover, primary blast samples from patients diagnosed with acute myeloid leukaemia recapitulated the senescence-conferred hypermetabolic phenotype and the susceptibility to autophagy-blocking or energy-depleting therapies (Fig. 4e–g). Hence, an energy-consuming endoplasmic reticulum



stress-autophagy cascade confers a critical liability to senescent cancer cells across species and entities.

Finally, we wondered whether such a liability may be therapeutically exploitable *in vivo*. Accordingly, exposure of CTX-senescent Eμ-*myc* control;*Bcl2* lymphomas to either 2DG or bafilomycin A1 *in vivo* resulted in rapid apoptotic tumour regression (Fig. 4h). Control;*Bcl2* lymphoma-bearing mice monitored for long-term outcome after CTX achieved a significantly extended overall survival when subsequently exposed to bafilomycin A1 treatment; all autophagy-co-targeted mice lived longer than their individually matched CTX-only counterparts (Fig. 4i and Supplementary Fig. 12; CTX-only cohort, median 51 days; CTX or bafilomycin A1 cohort, median 78 days;  $P = 0.038$ ; no significant differences were seen with bafilomycin A1 alone: median survival 26 days versus 24 days for observation only). Notably, no bafilomycin-A1-dependent improvement of the long-term outcome was detectable in cohorts of *Suv39h1*<sup>−</sup>;*Bcl2* lymphoma-bearing mice treated with CTX and bafilomycin A1 compared to mice treated with CTX only (Fig. 4i).

In essence, our data establish senescence-associated metabolic reprogramming as a cellular response to cope with SASP-coupled proteotoxic stress. High production of SASP factors (that is, protein synthesis, maturation, processing, vesicle transport and secretion) and senescence-related oxidative stress, further promoting the formation of misfolded or toxic proteins, account for the observed endoplasmic reticulum stress-UPR-ubiquitination autophagy cascade, the high energetic needs and the respective vulnerability of senescent cells (Supplementary Fig. 13). This cancer-specific vulnerability, which is not seen in normal cells that do not senesce in response to therapy (Supplementary Fig. 14), can be exploited selectively by new-concept, 'synthetic lethality'-like sequential treatment strategies, thereby adding to the clearance of senescent cells by the host immune system<sup>22–24</sup> (Supplementary Fig. 15). Although TIS is a desirable therapeutic outcome, particularly in apoptotically compromised tumours, the viable senescent cells have potentially harmful properties: they may be capable of cell-cycle reentry, affect stemness (M.M. and C.A.S., unpublished observations), secrete paracrine-active factors and promote inflammation. Therefore, subsequent elimination of senescent tumour cells should add considerably to the long-term efficacy of pro-senescent therapies. More generally, our findings illustrate how a cellular condition, not a distinct molecule, may serve as target for new-concept anticancer therapies.

## METHODS SUMMARY

Eμ-*myc* transgenic lymphomas with defined genetic defects were generated as described<sup>25</sup>. Methods used for quantitative real-time polymerase chain reaction (qRT-PCR), flow cytometry, immunostaining of cytospin preparations, growth-related assays, immunoblot analysis, fluorescence-based *in vivo* imaging, and statistical evaluation have been described previously<sup>3,25</sup>. See Methods for bioenergetics, metabolism or autophagy assays and related inhibitors, for bioluminescence imaging of metabolites *in situ*, and for PET and fluorescence imaging of whole animals.

**Online Content** Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 15 November 2011; accepted 4 July 2013.

Published online 14 August 2013.

1. Narita, M. *et al.* Rb-mediated heterochromatin formation and silencing of E2F target genes during cellular senescence. *Cell* **113**, 703–716 (2003).
2. Braig, M. *et al.* Oncogene-induced senescence as an initial barrier in lymphoma development. *Nature* **436**, 660–665 (2005).
3. Schmitt, C. A. *et al.* A senescence program controlled by p53 and p16<sup>INK4a</sup> contributes to the outcome of cancer therapy. *Cell* **109**, 335–346 (2002).

4. Acosta, J. C. *et al.* Chemokine signaling via the CXCR2 receptor reinforces senescence. *Cell* **133**, 1006–1018 (2008).
5. Coppe, J. P. *et al.* Senescence-associated secretory phenotypes reveal cell-nonautonomous functions of oncogenic RAS and the p53 tumor suppressor. *PLoS Biol.* **6**, e301 (2008).
6. Kuilman, T. *et al.* Oncogene-induced senescence relayed by an interleukin-dependent inflammatory network. *Cell* **133**, 1019–1031 (2008).
7. Campisi, J. & d'Adda di Fagnola, F. Cellular senescence: when bad things happen to good cells. *Nature Rev. Mol. Cell Biol.* **8**, 729–740 (2007).
8. Collado, M. & Serrano, M. Senescence in tumours: evidence from mice and humans. *Nature Rev. Cancer* **10**, 51–57 (2010).
9. Kuilman, T., Michaloglou, C., Mooi, W. J. & Peeper, D. S. The essence of senescence. *Genes Dev.* **24**, 2463–2479 (2010).
10. Nardella, C., Clohessy, J. G., Alimonti, A. & Pandolfi, P. P. Pro-senescence therapy for cancer treatment. *Nature Rev. Cancer* **11**, 503–511 (2011).
11. Dimri, G. P. *et al.* A biomarker that identifies senescent human cells in culture and in aging skin *in vivo*. *Proc. Natl Acad. Sci. USA* **92**, 9363–9367 (1995).
12. Moskowitz, C. H. *et al.* Risk-adapted dose-dense immunochemotherapy determined by interim FDG-PET in advanced-stage diffuse large B-cell lymphoma. *J. Clin. Oncol.* **28**, 1896–1903 (2010).
13. Jones, R. G. *et al.* AMP-activated protein kinase induces a p53-dependent metabolic checkpoint. *Mol. Cell* **18**, 283–293 (2005).
14. Warburg, O., Posener, K. & Negelein, E. Über den Stoffwechsel der Carcinomzelle. *Biochem. Z.* **152**, 319–344 (1924).
15. Vander Heiden, M. G. *et al.* Evidence for an alternative glycolytic pathway in rapidly proliferating cells. *Science* **329**, 1492–1499 (2010).
16. Young, A. R. *et al.* Autophagy mediates the mitotic senescence transition. *Genes Dev.* **23**, 798–803 (2009).
17. Chien, Y. *et al.* Control of the senescence-associated secretory phenotype by NF-κB promotes senescence and enhances chemosensitivity. *Genes Dev.* **25**, 2125–2136 (2011).
18. Jing, H. *et al.* Opposing roles of NF-κB in anti-cancer treatment outcome unveiled by cross-species investigations. *Genes Dev.* **25**, 2137–2146 (2011).
19. Kroemer, G., Marino, G. & Levine, B. Autophagy and the integrated stress response. *Mol. Cell* **40**, 280–293 (2010).
20. Pankiv, S. *et al.* p62/SQSTM1 binds directly to Atg8/lymphoma cells3 to facilitate degradation of ubiquitinated protein aggregates by autophagy. *J. Biol. Chem.* **282**, 24131–24145 (2007).
21. Nakagawa, T. *et al.* Caspase-12 mediates endoplasmic-reticulum-specific apoptosis and cytotoxicity by amyloid-beta. *Nature* **403**, 98–103 (2000).
22. Kaelin, W. G., Jr. The concept of synthetic lethality in the context of anticancer therapy. *Nature Rev. Cancer* **5**, 689–698 (2005).
23. Xue, W. *et al.* Senescence and tumour clearance is triggered by p53 restoration in murine liver carcinomas. *Nature* **445**, 656–660 (2007).
24. Kang, T. W. *et al.* Senescence surveillance of pre-malignant hepatocytes limits liver cancer development. *Nature* **479**, 547–551 (2011).
25. Reimann, M. *et al.* Tumor stroma-derived TGF-β limits Myc-driven lymphomagenesis via Suv39h1-dependent senescence. *Cancer Cell* **17**, 262–272 (2010).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank the late A. Harris, T. Jacks, T. Jenuwein, P. A. Khavari, N. Mizushima, D. Peeper, and M. Vander Heiden for mice, cells and materials; the flow cytometry facility at the Berlin-Brandenburg Center for Regenerative Therapies; N. Burbach, J. Dräger, A. Herrmann, K. Kirste, S. Maßwig, B. Teichmann and S. Spiesicke-Wegener for technical assistance; and members of the Schmitt laboratory for discussions and editorial advice. This work was supported by a Ph.D. fellowship to J.R.D. from the Boehringer Ingelheim Foundation, and grants from the Deutsche Forschungsgemeinschaft to W.M.-K. (MK576/15-1), to U.K. and A.K.B. (SFB 824), to U.K., B.D., S.L. and C.A.S. (SFB/TRR 54), and to C.A.S. from the Helmholtz Association (Helmholtz Alliance 'Preclinical Comprehensive Cancer Center'; grant no. HA-305) and the Deutsche Krebshilfe (grant no. 108789). This interdisciplinary work was made possible by the structural framework of the inter-institutional cooperation between Charité and MDC (now represented by the Berlin Institute of Health (BIH)), the Berlin School of Integrative Oncology (BSIO) graduate program funded within the Excellence Initiative, and the German Cancer Consortium (GCC).

**Author Contributions** J.R.D., S.L. and C.A.S. conceived the project, designed the experiments, and analysed the data, and W.M.-K., U.K., B.D., L.W. and St.K. provided critical input. Y.Y., G.B., C.Z., J.H.M.D., J.L., A.G., K.S., Su.K., S.W., M.G. and M.R. conducted experiments, M.M. compiled GSEA data, D.L. generated gene expression profiling data, M.H. analysed GEP data, B.P. carried out electron microscopy, and A.K.B. performed PET imaging. C.A.S., with editorial assistance from S.L., wrote the manuscript. All authors discussed the results and commented on the manuscript.

**Author Information** Microarray data are deposited at the Gene Expression Omnibus under accession numbers GSE31099 and GSE44355. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to C.A.S. ([clemens.schmitt@charite.de](mailto:clemens.schmitt@charite.de)).

## METHODS

**Mouse strains, lymphoma generation and primary AML samples.** All animal protocols used in this study were approved by the governmental review board (Landesamt Berlin), and conform to the appropriate regulatory standards. Lymphomas with defined genetic defects were generated in a C57BL/6 background by intercrossing Eμ-*myc* transgenic mice with mice carrying loss-of-function alleles at the *Suv39h1* locus<sup>25,26,27</sup>. *Suv39h1*<sup>−/−</sup> lymphomas reflect Eμ-*myc* lymphomas that arose in *Suv39h1*<sup>−/−</sup> females, or, owing to the X-linkage of the *Suv39h1* locus, in *Suv39h1*<sup>1Y/−</sup> males. Control and *Suv39h1*<sup>−/−</sup> lymphomas used for *in vitro* experiments were pre-screened for an intact p53 response by p53 and p21<sup>CIP1</sup> immunoblot analyses 6 h after γ-irradiation at a dose of 4 gray (Gy) (to exclude samples with a p53 pathway mutation; data not shown). Genotyping of the offspring by allele-specific genomic PCR, monitoring of lymphoma onset, preservation of snap-frozen or formalin-fixed lymph-node tissue and isolation of viable lymphoma cells (lymphoma cells) were carried out as described previously<sup>3,28</sup>. The use of small aliquots of human peripheral blood and bone marrow primarily obtained for the initial diagnosis of acute myeloid leukemia (AML) as anonymous samples was based on informed patient consent, and was specifically approved by the local ethics commission of the Charité – Universitätsmedizin Berlin (reference EA4/061/11). Mononuclear cells were isolated by Ficoll density-gradient centrifugation and red cell lysis, and cultured on feeder layers as described for murine lymphoma cells<sup>29</sup>.

**Lymphoma cell culture, retroviral gene transfer, and *in vitro* treatments.** Lymphoma cells or splenic B cells were cultured on irradiated NIH3T3 fibroblasts serving as feeder cells<sup>29</sup>. Retroviral gene transfer was carried out as described previously<sup>28</sup>; all retroviruses were murine stem cell virus (MSCV)-based constructs that either co-express puromycin (that is, MSCV-Bcl2-puro, MSCV-RasV12-puro) or GFP (that is, MSCV-Bcl2-IRES-GFP), including the corresponding vector-only controls (that is, MSCV-puro, or MSCV-IRES-GFP). For *in vitro* anticancer drug assays, adriamycin (ADR, Sigma), a topoisomerase II inhibitor, was added once in senescence-related experiments at a concentration of 0.01 μg ml<sup>−1</sup> for non-Bcl2-infected lymphoma cells and splenic B cells, and at a concentration of 0.05 μg ml<sup>−1</sup> for Bcl2-protected lymphoma cells. The serine/threonine-protein kinase mTOR inhibitor rapamycin was added once at a concentration of 5 μM to induce senescence. Stable knockdown of caspase 3 (shRNA: 5′-GGATCTATCTGGACAGTAGTT-3′) and caspase 12 (shRNA: 5′-GCTCTCATCATCTGCAA CAAA-3′) was achieved by retroviral infection of lymphoma cells with the pSuperRetro plasmid system (Oligo Engine). For lentiviral knockdown of *hexokinase 2* and *p62/SQSTM1*, individual complementary DNA constructs from the human pGIPz lentiviral shRNA library, version 13 (CSHL RNAmir, Open Biosystems) were used (hexokinase 2 lentiviral plasmid ID: V3LHS\_344623; *p62/SQSTM1* lentiviral plasmid ID: V2LMM\_77135). Viral particles were generated according to the manufacturer's instructions, using second generation helper plasmids (psPAX2, pMA2-VSVG) and the HEK293T packaging cell line.

***In vivo* treatments and mouse imaging.** Individual lymphomas were propagated by intravenous injection in one strain-matched, non-transgenic, fully immunocompetent mouse each and treated with a single intraperitoneal dose of cyclophosphamide (CTX, Sigma, 300 mg kg<sup>−1</sup> body weight) as described<sup>4</sup>. *In vivo*, CTX was preferred to ADR because of higher anti-lymphoma efficacy and less organ toxicity, whereas CTX itself is not active *in vitro*. In some experiments, mice were exposed to a single dose of the following drugs 5 days after CTX treatment: 2DG (Sigma, 250 mg kg<sup>−1</sup> body weight intraperitoneally), bafilomycin A1 (Sigma, 1 μg in 150 μl phosphate buffer intraperitoneally), or phosphate buffer as mock treatment. Treatment responses *in vivo* (time to death, reflecting 'overall survival') were monitored as described previously<sup>29</sup>. After CO<sub>2</sub> euthanasia, lymph nodes were isolated to generate tissue sections or single-cell suspensions as described previously<sup>29</sup>. For *in vivo* imaging, mice were anaesthetized with isoflurane (Abbott) using a veterinary anaesthesia system (Vetlad Medical Sales and Services, and Harvard Apparatus). Fluorescence-based whole-body imaging was carried out with a charge-coupled device (CCD)-camera-equipped bioimaging system (LAS-4000, Fujifilm).

**FDG- and FLT-PET imaging.** The glucose analogue FDG was produced at the cyclotron unit of the Technische Universität in Munich using a standard procedure according to good manufacturing practice (GMP) guidelines. The nucleoside FLT was produced by a method described previously<sup>30</sup> with minor modifications using an automated radiotracer synthesizer (General Electric Medical Systems). FDG- and FLT-PET were performed using a dedicated small animal PET system (INVEON, Siemens Preclinical Solutions) within 24 h before treatment, and early in the course of therapy (typically on day 5 or 6). Static imaging was started 45 min after intravenous injection of 3.7–7.4 megabecquerels (MBq) FLT or FDG per mouse. Data were acquired over 15 min, normalized and corrected for random coincidence events, dead time, and decay. To assess tracer accumulation at tumour sites, tumour-to-background ratios (TBRs) were calculated for all tumour manifestations that showed focal tracer accumulation prior to therapy. Circular two-dimensional

regions of interest (ROIs), not covering the entire tumour to avoid partial volume effects, were placed manually in the area with the highest tumour activity. To determine background activity, a corresponding ROI was set at the muscle of the right thigh. Corresponding TBR values were calculated as mean count rates within the tumour and the matching background ROI.

### Cell growth, apoptosis, autophagy and cellular senescence *in vitro* and *in vivo*.

Viability and cell numbers were analysed by trypan blue dye exclusion. DNA synthesis was assessed by immunohistochemical staining of incorporated BrdU as described previously<sup>25,28</sup>. Detection of cell proliferation by Ki67 staining and assessment of SA-β-gal activity at pH 5.5 (mouse) or pH 6.0 (human) in cryosections or cytospin preparations of cell suspensions were carried out as described previously<sup>2,3</sup>. Apoptotic DNA strand breaks were detected by TUNEL staining in paraffin-embedded tissue sections in accordance with the manufacturer's (Roche) protocol<sup>2,3</sup>.

**Global protein synthesis and gene-expression analyses.** Total protein synthesis was detected by the Click-iT AHA Alexa Fluor 488 Protein Synthesis HCS Assay (Invitrogen) according to the manufacturer's protocol. Cells were incubated with medium containing Click-iT-labelled L-methionine for 30 min, washed and subsequently loaded with Click-iT-reactive Alexa Fluor 488 dye. DAPI was used as a counterstain. Images were automatically acquired using a Pathway 855 high-content bioimager confocal microscope (Becton Dickinson) and subsequently analysed using the AttoVision software package (version 1.6) as a threshold-based single-cell densitometric measurement of the Alexa Fluor 488 signal.

Genomic DNA was isolated as described<sup>29</sup>. For RT-PCR analyses, RNA extracts isolated with Trizol (Invitrogen) were transcribed into cDNA using SuperScript reverse transcriptase (Invitrogen) and random hexamers or oligo-dT. Primer sequences and detailed PCR protocols for the detection of the Eμ-*myc* transgene and α-tubulin (as an internal control) as well as for quantitative RT-PCR (qRT-PCR) analyses of murine *Gapdh*, *Casp3*, *Casp12*, *Cpt1a*, *Cxcl5*, *Glut1*, *Glut3*, *Hk2*, *Igf1bp7*, *Il-6*, *Mcp1*, *Mmp2*, *p62/SQSTM1*, *Serpine1* (also known as *Pai-1*), *Pfk1*, and *Sco2* as well as human *GAPDH*, *CXCL1*, *HGF*, *IL-6*, *IL-8*, and *MMP3* (using commercially available primers; Applied Biosystems) are available on request. For every sample, the difference in cycle threshold (ΔCt) values were determined as the difference between the Ct value of a specific transcript and the Ct value of *GAPDH*, serving as the housekeeping control messenger RNA, and relative transcript levels (for example, treated versus untreated) were calculated based on 2<sup>(−ΔΔCt)</sup> with ΔΔCt = ΔCt<sub>treated</sub> − ΔCt<sub>untreated</sub>. Detection of *XBPI* gene splicing was carried out as described previously<sup>31</sup>.

Examination of proteins by immunofluorescence and immunoblot analysis was carried out as reported previously<sup>32</sup>. Whole-cell extracts for immunoblot analysis were generated by lysing cells in NP-40 protein lysis buffer (150 mM NaCl, 50 mM Tris-HCl, 1% NP-40, pH 8.0) including proteinase inhibitors and phosphatase inhibitors. Equal aliquots corresponding to 30 to 60 μg of protein were resolved on SDS-polyacrylamide gels and transferred to Immobilon-P membranes (Millipore). Antigen detection was carried out with antibodies against AMPK (no. 2603, Cell Signaling Technology (CST), 1:1000), AMPK-P-T172 (no. 2535, CST, 1:1000), ATF4 (c20, Santa-Cruz Biotechnology (SCBT), 1:250), BrdU (Molecular Probes, 1:2000), Bcl2 (SCBT, 1:500), CHOP (sc-7351, Santa Cruz, 1:200), Caspase 4 (no. 4450, CST, 1:1000), Caspase 12 (no. 2202, CST, 1:1000), cleaved (D175) Caspase 3 (no. 9661, CST, 1:1000), JNK (no. 9258, CST, 1:1000), JNK-P-T183/T185 (no. 4671, CST, 1:1000), p62/SQSTM1 (no. 5114, CST, 1:1000), PKM1 and PKM2 (both gifts from M. Vander Heiden; 1:1000), microtubule-associated protein 1 light chain 3 (2G6, Nanotools, 1:1000), α-tubulin (T5168, Sigma-Aldrich, 1:8000), ubiquitin (no. 3933, CST, 1:1000), and V-ATPase A1 (sc28801, Santa Cruz, 1:250). Where required, corresponding anti-mouse and anti-rabbit peroxidase-conjugated secondary antibodies (GE Healthcare; 1:1000 and 1:2000, respectively) were used.

For microarray-based gene-expression profiling, RNA was isolated from lymphoma cells using the RNeasy Mini kit (Qiagen), and hybridized to Affymetrix Mouse Genome 430 2.0 microarrays (Affymetrix) according to the manufacturer's instructions. Arrays were hybridized, washed and scanned by standard Affymetrix protocols. The resulting CEL files were imported into Partek Genomic Suite Software (Version 6.4; Partek), and the data were processed by the implemented default workflow. The raw microarray data were deposited at the Gene Expression Omnibus (GEO) repository of the National Center for Biotechnology Information under accession numbers GSE31099 and GSE44355.

The enrichment of distinct, Gene-Ontology-defined gene sets within the global transcript signals from matched pairs of untreated versus ADR-treated control; *Bcl2* or *Suv39h1*<sup>−/−</sup>*Bcl2* lymphoma cells was analysed by applying the GSEA v2.0 software (Broad Institute of the Massachusetts Institute of Technology (MIT) and Harvard, <http://www.broad.mit.edu/gsea>). Normalized enrichment scores (NES) reflect a statistically significant enrichment for *P* values below 0.05 and false discovery rate (FDR) values below 0.25 (ref. 33).



**Analysis of cell metabolism.** *In situ* analyses of glucose levels in cryopreserved lymphoma tissue sections were performed by bioluminescence imaging as described<sup>34</sup>. Glucose was enzymatically linked to the luciferase enzyme to obtain a light emission signal whose intensity was proportional to the tissue glucose concentration. The spatial distribution of the bioluminescence intensity was registered directly using a microscope (Axiophot, Zeiss) and an electron multiplying charge coupled device (EMCCD) low-light imaging camera (iXon, Andor Technology). The light intensity was calibrated as weight-related tissue concentrations ( $\mu\text{mol g}^{-1}$ ) by appropriate tissue standards.

For stable isotope-resolved metabolomics (SIRM), lymphoma cells were pooled and centrifuged at 300g for 5 min. Subsequently, cell pellets were resuspended in pre-warmed labelling medium containing  $^{13}\text{C}_6$ -glucose ( $2.5 \text{ g l}^{-1}$ , Sigma) as a labelled carbon source, and incubated for 3 min at  $37^\circ\text{C}$ . Cell suspensions were centrifuged (3,000g for 5 min at  $4^\circ\text{C}$ ) and the supernatants discarded. Cell pellets were quenched in liquid nitrogen to stop all metabolic processes. Intracellular metabolites were prepared by methanol–chloroform–water extraction. For extracellular metabolites, lymphoma cells were propagated in 24-well plates, and 200  $\mu\text{l}$  of medium was collected from independent wells at time-points as indicated, and stored at  $-25^\circ\text{C}$  until sample extraction. Fifty microlitres of each media sample were extracted with 1 ml pre-chilled methanol–chloroform–water buffer (5:2:1 v/v) under agitation for 30 min at  $4^\circ\text{C}$ . Water (0.5 ml) was then added to the solution (intra- or extracellular metabolites). The upper polar phase was dried and stored at  $-20^\circ\text{C}$  for subsequent analyses.

For gas chromatography–mass spectrometry (GC-MS) analysis of intracellular and extracellular metabolites, cell pellets were prepared and measured as described<sup>35</sup>. The GC-MS chromatograms were processed with the ChromaTOF software (LECO). Data matrices for the relative quantification were extracted from the mass spectra using the MetMax software<sup>35,36</sup>.

*In vitro* analyses of metabolites were carried out by enzyme-linked absorbance or luminescence kits for glucose (Amplex Red Glucose/Glucose Oxidase kit, Invitrogen), lactate (Lactate assay kit, BioVision) and ATP (APOSensor ATP luminescence assay kit, Biovision); or, in some experiments, ATP colorimetric/fluorometric assay kit, Biovision). Glucose uptake was assessed by flow cytometry after exposure to  $10 \mu\text{M}$  of the fluorescent glucose analogue 2-(N-(7-nitrobenz-2-oxa-1,3-diazol-4-yl)amino)-2-deoxyglucose (2NBDG; Molecular Probes) for 2 h. OCRs and extracellular acidification rates (ECARs) were measured in an extracellular flux analyser in accordance with the manufacturer's standard protocol (XF24, Seahorse Biosciences). Glucose metabolism was blocked by glucose transport inhibitors phloretin (Sigma,  $50 \mu\text{M}$ ) or cytochalasin B (Sigma,  $10 \mu\text{M}$ ) or by 2-deoxy-D-glucose (Sigma,  $5 \text{ mM}$ ). Glucose-free medium was supplemented with D-Glucose (Sigma,  $5 \text{ mM}$ ). The lactate dehydrogenase enzyme was inhibited by sodium oxamate (Sigma,  $25 \text{ mM}$ ). The electron transport chain was poisoned with antimycin A (Sigma,  $1 \mu\text{M}$ ). AMPK was inhibited by Compound C (Sigma,  $10 \mu\text{M}$ ). Fatty acid oxidation was blocked by the carnitine palmitoyltransferase I inhibitor etomoxir (Sigma,  $100 \mu\text{M}$ ). Lysosomal V-ATPases were inhibited by bafilomycin A1 (Sigma,  $5 \mu\text{M}$ ) or concanamycin A (Sigma,  $50 \text{ nM}$ ), and lysosomal proteases were blocked using a cocktail of inhibitors (composed of pepstatin A, E64d, and ZPAD (N-CBZ-L-phenylalanyl-L-alanine-diazomethylketone), all from Sigma,  $10 \mu\text{M}$ ). Inhibitors were added once 48 h before analysis.

For global metabolite analyses, lymphoma cells were subjected to a methanol–chloroform–water extraction as reported<sup>37</sup>. The polar fraction ( $150 \mu\text{l}$ ) was dried, derivatized and analysed by GC-MS as described<sup>38</sup>. The R software package TargetSearch was used to compare measured data against a reference library (Golm Metabolome Database, <http://gmd.mpimp-golm.mpg.de>), and to extract the intensity of quantification ions for 109 metabolites<sup>39</sup>. The non-polar fraction ( $500 \mu\text{l}$ ) was analysed using an Exactive mass spectrometer (Thermo-Fisher). The intensity of 140 representative ion masses was retained after filtering for non-redundancy, abundance, assignable sum formulas and matching isotopic peaks using custom-made R scripts. Both raw data matrices were  $\log_{10}$  transformed to improve normality and corrected for run-order effects and differences in sample median intensity using an appropriate analysis of variance (ANOVA) model<sup>40</sup>. For statistical analyses, normalized metabolite data were subjected to a principal component analysis (PCA) using the nipsals algorithm from the pcaMethods package<sup>41</sup> available for the statistical software framework R (<http://www.r-project.org/>) following an initial Pareto scaling.

Fatty acid oxidation was assessed by quantification of 16:0 phosphatidylcholine and its oxidized form lyso-phosphatidylcholine. Measurements of lymphoma cell lysates were performed according to an established protocol using liquid chromatography coupled to triple-quadrupole mass spectrometry (LC-MS-MS)<sup>42</sup>. Multiple reaction monitoring transitions were as follows: phosphatidylcholine (16:0 18:2)  $m/z$  758/184, lyso-phosphatidylcholine (16:0)  $m/z$  496/184. Liquid chromatography resolution of all analytes was achieved using a  $2 \times 60 \text{ mm}$  MultoHigh C18 reversed phase column with  $3 \mu\text{m}$  particle size (CS-Chromatographie Service). Linearity of

the standard curves and correlation coefficients were obtained by linear regression analyses. Data analyses were performed using Analyst 1.4 (ABSciex).

**Oncogene-induced senescence.** Wild-type,  $p53^{-/-}$  and  $atg5^{-/-}$  mouse embryo fibroblasts (MEFs) were produced in accordance with regulatory standards of the governmental review board (Landesamt Berlin). MEFs and the human diploid fibroblast cell line Tig3 (a gift from D. Peeper) were cultured in accordance with published protocols<sup>43</sup>. OIS was established by stable transduction of cells with an MSCV-RasV12-blasticidin retrovirus, or infection with a 4-hydroxy-tamoxifen (4-OHT)-inducible MSCV-RasV12•endoplasmic reticulum-blasticidin retrovirus, encoding an oncogenic Ras-to-oestrogen receptor (ER) fusion (RasV12•ER, a gift from P. A. Khavari), and the addition of  $1 \mu\text{M}$  of 4-OHT. In some experiments, Tig3 cells were infected with the pSuperRetro plasmid system for stable knockdown of  $p53$  (shp53) or *IL-6* (shIL-6)<sup>44</sup>. Senescence was detected by SA- $\beta$ -gal staining at day 5 after successful antibiotic selection or 4-OHT addition.

**Autophagosome, endoplasmic reticulum and lysosome structure.** Cellular ultrastructure and formation of autophagosomes were assessed by electron microscopy. Lymphoma cells were fixed in a 0.1 M phosphate buffer containing 4% formaldehyde and 1% glutaraldehyde for 24 h. Samples were stained with 1%  $\text{OsO}_4$  for 2 h, dehydrated in a graded ethanol series and propylene oxide and embedded in Poly/Bed 812 (Polysciences). Ultrathin sections were contrasted with uranyl acetate and lead citrate. For cryosectioning, cells were fixed in a 0.1 M phosphate buffer containing 4% formaldehyde and 0.5% glutaraldehyde for 1 h, subsequently embedded in 10% gelatine, and infiltrated with 2.3 M sucrose. Ultrathin cryosections were contrasted and stabilized with a mixture of 3% tungstosilicic acid hydrate (Fluka) and 2.5% polyvinyl alcohol (Sigma) as described<sup>45</sup>. Sections were examined with a transmission electron microscope (Morgagni, FEI). Digital images were taken with a CCD camera (Morada), equipped with the iTEM software package (Olympus Soft Imaging Solutions).

Lysosomes were stained with LysoTracker Red DND-99 (Molecular Probes) as recommended by the manufacturer's protocol. Lymphoma cells were incubated with  $100 \text{ nM}$  LysoTracker for 1 h in cell culture medium, then attached to cover slips and imaged by fluorescence microscopy.

**Intracellular ATP levels.** Control; *Bcl2* and *Suv39h1*<sup>-/-</sup>; *Bcl2* lymphoma cells were subjected to a potentially senescence-inducing ADR treatment for 5 days, and subsequently seeded in 384-well plates, where they were exposed to  $100 \text{ nM}$  Concanamycin A for 10 min or left untreated. ATP levels were measured by an ATP luminescent assay kit (Promega) according to the manufacturer's instructions.

**Calcium efflux.** Intracellular  $\text{Ca}^{2+}$  efflux was measured in accordance to the manufacturer's protocol (Fluo-4 NW Assay, Invitrogen). Cells were resuspended in assay buffer, and seeded onto a poly-L-lysine-coated 96-well imaging plate. After a 1-hour incubation at  $37^\circ\text{C}$ , the calcium-sensitive dye Fluo-4 was added to each well. Following a 30-min equilibration period of the plate to obtain a stable baseline intensity, 2-deoxy-D-glucose or bafilomycin A1 were injected and cells were automatically analyzed every 20 to 40 s. Fluorescence intensity curves were normalized against the values obtained from untreated cells.

**Human cell lines.** The human B-cell lymphoma (Mec1, RCK8), colon carcinoma (CaCo2, HCT116), melanoma (A-375, MEWO, SkMel28), myeloid leukemia (KG1, K562), fibrosarcoma (HT1080) and non-small cell lung cancer (A549) cell lines, obtained from local partner laboratories, were exposed to ADR (lymphomas,  $10 \text{ ng ml}^{-1}$ ; leukaemias,  $25 \text{ ng ml}^{-1}$ ; colon carcinomas and melanomas,  $50 \text{ ng ml}^{-1}$ ; lung carcinoma,  $100 \text{ ng ml}^{-1}$ ) to induce TIS. Six days after initiation of treatment, cells were analysed for SA- $\beta$ -gal activity and subjected to additional experiments. Some of the cell lines (A375, Mec1, RCK8) were lately authenticated using multiplex cell authentication (Multiplexion)<sup>46</sup>. The SNP profiles matched known profiles or were unique. Moreover, no mycoplasma contamination was detected in regular screenings of our cell lines.

**Statistics.** Based on previous experience with the *Eμ-myc* transgenic mouse lymphoma model, sample sizes typically reflect three to five individual primary tumours (that is, biological, not technical replicates), or even ten or more tumour-bearing animals per arm in investigator-blinded survival analyses of long-term outcome to treatment *in vivo*. If not stated otherwise, data are presented as arithmetic means  $\pm$  s.d. or s.e.m., and statistical analyses of the generally normally distributed (as assessed by the Shapiro–Wilk test) and variance-homogeneous (as assessed by the Levene test) data were based on paired or unpaired *t*-tests. The survival of mouse subcohorts was calculated using Kaplan–Meier estimation, and compared with the log-rank (Mantel–Cox) test. For GSEA, the non-parametric Kolmogorov–Smirnov test was applied. All quantifications from staining reactions were carried out by an independent and blinded second examiner, and reflect at least three samples with at least 200 events counted (typically in three different areas) each. Unless otherwise stated, a *P* value of  $<0.05$  was considered statistically significant.

26. Adams, J. M. *et al.* The *c-myc* oncogene driven by immunoglobulin enhancers induces lymphoid malignancy in transgenic mice. *Nature* **318**, 533–538 (1985).



27. Peters, A. H. *et al.* Loss of the *Suv39h* histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* **107**, 323–337 (2001).
28. Schmitt, C. A. *et al.* Dissecting p53 tumor suppressor functions *in vivo*. *Cancer Cell* **1**, 289–298 (2002).
29. Schmitt, C. A., McCurrach, M. E., de Stanchina, E., Wallace-Brodeur, R. R. & Lowe, S. W. *INK4a/ARF* mutations accelerate lymphomagenesis and promote chemoresistance by disabling p53. *Genes Dev.* **13**, 2670–2677 (1999).
30. Shields, A. F. *et al.* Imaging proliferation *in vivo* with [F-18]FLT and positron emission tomography. *Nature Med.* **4**, 1334–1336 (1998).
31. Marciniak, S. J. *et al.* CHOP induces death by promoting protein synthesis and oxidation in the stressed endoplasmic reticulum. *Genes Dev.* **18**, 3066–3077 (2004).
32. Reimann, M. *et al.* The Myc-evoked DNA damage response accounts for treatment resistance in primary lymphomas *in vivo*. *Blood* **110**, 2996–3004 (2007).
33. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
34. Walenta, S. *et al.* High lactate levels predict likelihood of metastases, tumor recurrence, and restricted patient survival in human cervical cancers. *Cancer Res.* **60**, 916–921 (2000).
35. Liu, L. *et al.* Deregulated MYC expression induces dependence upon AMPK-related kinase 5. *Nature* **483**, 608–612 (2012).
36. Kempa, S. *et al.* An automated GCxGC-TOF-MS protocol for batch-wise extraction and alignment of mass isotopomer matrixes from differential <sup>13</sup>C-labelling experiments: a case study for photoautotrophic-mixotrophic grown *Chlamydomonas reinhardtii* cells. *J. Basic Microbiol.* **49**, 82–91 (2009).
37. Giavalisco, P. *et al.* High-resolution direct infusion-based mass spectrometry in combination with whole <sup>13</sup>C metabolome isotope labeling allows unambiguous assignment of chemical sum formulas. *Anal. Chem.* **80**, 9417–9425 (2008).
38. Lisec, J., Schauer, N., Kopka, J., Willmitzer, L. & Fernie, A. R. Gas chromatography mass spectrometry-based metabolite profiling in plants. *Nature Protocols* **1**, 387–396 (2006).
39. Cuadros-Inostroza, A. *et al.* TargetSearch—a Bioconductor package for the efficient preprocessing of GC-MS metabolite profiling data. *BMC Bioinformatics* **10**, 428 (2009).
40. Lisec, J. *et al.* Corn hybrids display lower metabolite variability and complex metabolite inheritance patterns. *Plant J.* **68**, 326–336 (2011).
41. Stacklies, W., Redestig, H., Scholz, M., Walther, D. & Selbig, J. *pcaMethods*—a bioconductor package providing PCA methods for incomplete data. *Bioinformatics* **23**, 1164–1167 (2007).
42. Bode, C. & Graler, M. H. Quantification of sphingosine-1-phosphate and related sphingolipids by liquid chromatography coupled to tandem mass spectrometry. *Methods Mol. Biol.* **874**, 33–44 (2012).
43. Serrano, M., Lin, A. W., McCurrach, M. E., Beach, D. & Lowe, S. W. Oncogenic *ras* provokes premature cell senescence associated with accumulation of p53 and p16<sup>INK4a</sup>. *Cell* **88**, 593–602 (1997).
44. Berns, K. *et al.* A large-scale RNAi screen in human cells identifies new components of the p53 pathway. *Nature* **428**, 431–437 (2004).
45. Reimer, T. A. *et al.* Reevaluation of the 22-1-1 antibody and its putative antigen, EBAG9/RCAS1, as a tumor marker. *BMC Cancer* **17**, 47 (2005).
46. Castro, F. *et al.* High-throughput SNP-based authentication of human cell lines. *Int. J. Cancer* **132**, 308–314 (2013).

# Bacterial colonization factors control specificity and stability of the gut microbiota

S. Melanie Lee<sup>1</sup>, Gregory P. Donaldson<sup>1\*</sup>, Zbigniew Mikulski<sup>2\*</sup>, Silva Boyajian<sup>1</sup>, Klaus Ley<sup>2</sup> & Sarkis K. Mazmanian<sup>1</sup>

Mammals harbour a complex gut microbiome, comprising bacteria that confer immunological, metabolic and neurological benefits<sup>1</sup>. Despite advances in sequence-based microbial profiling and myriad studies defining microbiome composition during health and disease, little is known about the molecular processes used by symbiotic bacteria to stably colonize the gastrointestinal tract. We sought to define how mammals assemble and maintain the *Bacteroides*, one of the most numerically prominent genera of the human microbiome. Here we find that, whereas the gut normally contains hundreds of bacterial species<sup>2,3</sup>, germ-free mice mono-associated with a single *Bacteroides* species are resistant to colonization by the same, but not different, species. To identify bacterial mechanisms for species-specific saturable colonization, we devised an *in vivo* genetic screen and discovered a unique class of polysaccharide utilization loci that is conserved among intestinal *Bacteroides*. We named this genetic locus the commensal colonization factors (*ccf*). Deletion of the *ccf* genes in the model symbiont, *Bacteroides fragilis*, results in colonization defects in mice and reduced horizontal transmission. The *ccf* genes of *B. fragilis* are upregulated during gut colonization, preferentially at the colonic surface. When we visualize microbial biogeography within the colon, *B. fragilis* penetrates the colonic mucus and resides deep within crypt channels, whereas *ccf* mutants are defective in crypt association. Notably, the CCF system is required for *B. fragilis* colonization following microbiome disruption with *Citrobacter rodentium* infection or antibiotic treatment, suggesting that the niche within colonic crypts represents a reservoir for bacteria to maintain long-term colonization. These findings reveal that intestinal *Bacteroides* have evolved species-specific physical interactions with the host that mediate stable and resilient gut colonization, and the CCF system represents a novel molecular mechanism for symbiosis.

International microbiome sequencing initiatives are revealing detailed inventories of diverse bacterial communities across various body sites, diets and human populations<sup>2–4</sup>. Complex ecosystems have been forged by co-adaptation over millennia between animals and microbes to create stable and specific microbiomes<sup>5,6</sup>, suggesting the evolution of molecular mechanisms that establish and maintain symbiotic microbial colonization. Bacteroidetes is one of the most numerically abundant Gram-negative phyla in the mammalian gastrointestinal tract<sup>7</sup>. Studies in the genus *Bacteroides* have revealed species that induce glycosylation of the intestinal epithelium<sup>8</sup>, produce glycoside hydrolases that digest carbohydrates for host nutrient use<sup>9</sup>, direct host immune maturation<sup>10</sup> and protect animals from inflammation in experimental models of inflammatory bowel disease and multiple sclerosis<sup>11–13</sup>. To explore the dynamics of microbiome assembly, we sequentially introduced *Bacteroides* species to germ-free mice and monitored colonization via colony-forming units (c.f.u.) in faeces. Animals were readily colonized with *B. fragilis* followed by *Bacteroides thetaiotaomicron* (Fig. 1a) or *Bacteroides vulgatus* (Fig. 1b), and altering the sequence of microbial exposure does not affect results (Supplementary Fig. 1a). Notably, however, animals colonized with *B. fragilis* and then exposed

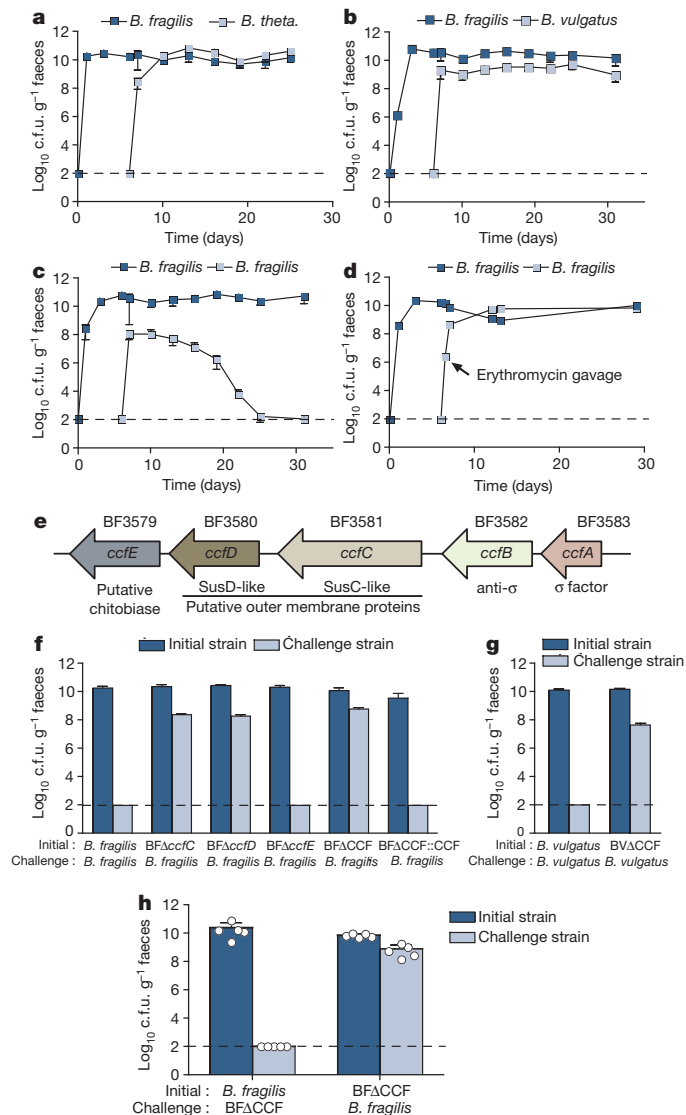
to the same species (marked by an antibiotic resistance gene) are resistant to super-colonization and clear the challenging strain (Fig. 1c). This novel observation of ‘colonization resistance’ by the same species is conserved in three other *Bacteroides* (Supplementary Fig. 1b–d), regardless of the antibiotic resistance markers used (Supplementary Fig. 1e), but not in *Escherichia coli* (Supplementary Fig. 1f). As conventional mice typically harbour 10<sup>11</sup>–10<sup>12</sup> c.f.u. per gram of caecal content<sup>14</sup> (100-fold greater than *Bacteroides* in mono-association), there does not seem to be a shortage of space or nutrients under these conditions, using a nutrient-rich standard diet. We thus proposed that individual *Bacteroides* species colonize the gut by saturating a limited and unique niche. Indeed, treatment of *B. fragilis* mono-associated mice with erythromycin to displace the existing strain permits colonization by an erythromycin-resistant challenge strain (Fig. 1d). These data suggest that *Bacteroides* colonize the gut in a species-specific and saturable manner.

We developed a functional *in vivo* screen to identify genetic factor(s) from *B. fragilis* that are sufficient to mediate species-specific colonization. Mice were mono-associated with *B. vulgatus*, then challenged with a library of *B. vulgatus* clones that each contained a fragment of *B. fragilis* genomic DNA (schematic in Supplementary Fig. 2a). We reasoned that only those clones containing genes that conferred stable gut colonization by *B. fragilis* would persist, with the remainder being cleared. We screened 2,100 clones each containing 9–10 kilobases of DNA, providing a 3.8-fold coverage of the *B. fragilis* genome and 98% probability that a given DNA sequence is present in the library (Supplementary Equation (1)). Notably, 30 days after orally gavaging the library into animals, only two clones sustained colonization. The inserts from both clones mapped to the same locus on the *B. fragilis* genome, BF3579–BF3583 (Supplementary Fig. 2b).

On the basis of predicted protein sequences, BF3583 and BF3582 constitute a sigma ( $\sigma$ ) factor/anti- $\sigma$  factor gene pair. BF3581 is a member of the SusC family of outer membrane proteins. BF3580 is a homologue of SusD, a lipoprotein often paired with SusC. These Sus-like systems have been shown to bind and import a range of oligosaccharide molecules<sup>15–18</sup>. BF3579 encodes a putative chitinase, suggesting a possible polysaccharide substrate for this system<sup>19</sup> (Fig. 1e). Comparative genomic analysis using the Integrated Microbial Genomes database (<http://img.jgi.doe.gov/cgi-bin/w/main.cgi>) reveals conservation of similar clusters of genes among sequenced intestinal *Bacteroides* species (Supplementary Fig. 3). Sus-like systems are numerous in *Bacteroides* within polysaccharide utilization loci (PULs), which are gene cassettes used to harvest dietary sugars and/or forage host glycans during nutrient deprivation<sup>15,20,21</sup>. As PULs have not previously been implicated in saturable niche colonization, the locus we have identified encodes a unique pathway in *Bacteroides* for species-specific gut association; we named the genes *ccfA–E*, for commensal colonization factors (Fig. 1e). Furthermore, as deletion of the most closely related genes from *B. fragilis* (BFΔ0227–0229; Supplementary Fig. 3) do not affect colonization dynamics (Supplementary Fig. 4) we suggest that the CCF system represents a functionally unique subset of PULs that evolved to promote long-term symbiosis.

<sup>1</sup>Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, California 91125, USA. <sup>2</sup>Division of Inflammation Biology, La Jolla Institute for Allergy and Immunology, La Jolla, California 92037, USA.

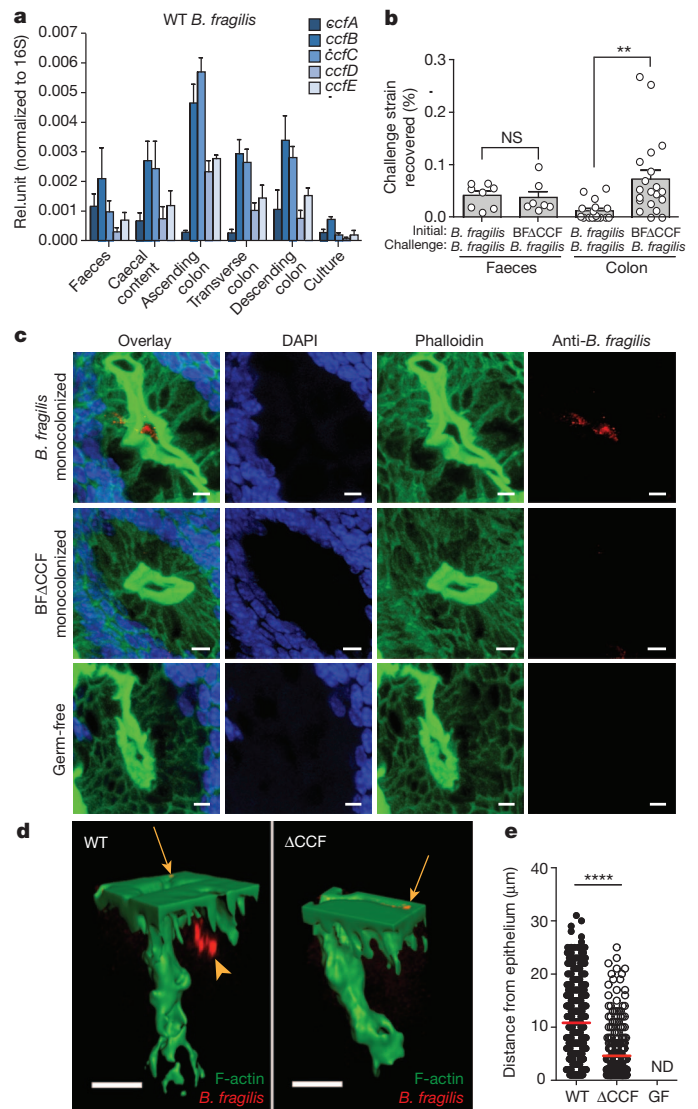
\*These authors contributed equally to this work.



**Figure 1** | *Bacteroides* species occupy species-specific niches in the gut via an evolutionarily conserved genetic locus. **a–c**, Germ-free mice were mono-associated with *B. fragilis* and challenged orally with *B. theta* (a), *B. vulgatus* (b) or *B. fragilis* (c). **d**, Mice were mono-associated with erythromycin-sensitive *B. fragilis*, and subsequently challenged with erythromycin-resistant *B. fragilis*. Erythromycin was administered where indicated. **e**, Genomic organization of the *ccf* locus. **f**, Mice were mono-associated with either wild-type *B. fragilis*, mutant strains deleted in *ccfC*, *ccfD*, *ccfE*, and *ccfC–E* (*BF $\Delta$ CCF*), or complemented strain (*BF $\Delta$ CCF::CCF*) and challenged with wild-type *B. fragilis*. c.f.u. were determined after 30 days. **g**, Mice were mono-associated with wild-type *B. vulgatus* or a mutant strain deleted in *ccfC–E* genes (*BV $\Delta$ CCF*), and challenged with wild-type *B. vulgatus*. c.f.u. were determined after 30 days. In all sequential colonization studies, results are representative of at least two independent trials ( $n = 3–4$  animals per group). **h**, Cross-colonization between wild-type *B. fragilis* and *BF $\Delta$ CCF* mono-associated mice at 7 days after encounter measured by c.f.u. of the initially colonizing and the horizontally transmitted (challenge) strains ( $n = 2$  animals per encounter, 5 independent trials). All graphs: dashed line indicates the limit of detection at  $100$  c.f.u.  $g^{-1}$  faeces, and error bars indicate s.d.

To test whether the putative structural genes (*ccfC–E*) are required for gut colonization, we generated in-frame deletion mutants of *B. fragilis*:  $\Delta$ *ccfC*,  $\Delta$ *ccfD* and  $\Delta$ *ccfE*. All strains exhibit normal morphology on solid agar medium and unimpaired growth in laboratory culture (data not shown). As shown previously, animals mono-colonized with wild-type *B. fragilis* completely clear the wild-type challenge strain after 30 days (Fig. 1f; first bars). However, animals mono-associated with  $\Delta$ *ccfC* or  $\Delta$ *ccfD* are permissive to colonization by wild-type bacteria (Fig. 1f;

second and third bars), unlike the *ccfE* mutant (Fig. 1f; fourth bars). A deletion mutant in all three genes (*BF $\Delta$ CCF*) also allows wild-type *B. fragilis* to colonize (Fig. 1f; fifth bars). Trans-complementation of the *BF $\Delta$ CCF* strain with *ccfA–E* restores colonization resistance (Fig. 1f; sixth bars). Similarly, a mutant in the *B. vulgatus* *ccfC–E* orthologues ( $\Delta$ BVU0946–BVU0948) also permits wild-type *B. vulgatus* to colonize (Fig. 1g), demonstrating conservation in *Bacteroides* species. When we tested horizontal transmission of bacteria between wild-type *B. fragilis*



**Figure 2** | *B. fragilis* colonization of the colonic crypts is mediated by the CCF system. **a**, Quantitative PCR with reverse transcription (qRT-PCR) of *ccf* gene expression levels normalized to 16S rRNA ( $n = 3$  animals, 2 trials). **b**, Mice were mono-associated with either wild-type *B. fragilis* or *BF $\Delta$ CCF*, and challenged with wild-type *B. fragilis*. The percentage of challenge strain was determined in the lumen (faeces) and colon after 1 day ( $n = 8$  animals per group). **c**, Confocal micrographs of germ-free, wild-type *B. fragilis* or *BF $\Delta$ CCF* mono-associated mice colon whole-mount. Crypts are visualized by DAPI (4',6-diamidino-2-phenylindole; nuclei, blue) and phalloidin (F-actin, green). Bacteria (red) are stained with IgY polyclonal antibody raised against *B. fragilis*. Images are representative of seven different sites analysed from at least two different colons. Scale bars,  $5 \mu m$ . **d**, Three-dimensional reconstructions of colon crypts from wild-type *B. fragilis* or *BF $\Delta$ CCF* mono-associated mice. Bacteria are detected on the apical surface of the epithelium (arrows) and in the crypt space (arrowhead). Scale bars,  $10 \mu m$ . **e**, Quantification of bacterial penetration, measured as distance from the epithelial surface per crypt. Error bars indicate s.e.m. GF, germ-free; ND, not detected; NS, not significant; rel.unit, relative unit; WT, wild-type. \*\*\*\*  $P < 0.0001$ .



and BF $\Delta$ CCF mono-associated mice, only wild-type bacteria cross-colonized (Fig. 1h). Thus, the CCF system is involved in colonization resistance by *Bacteroides*.

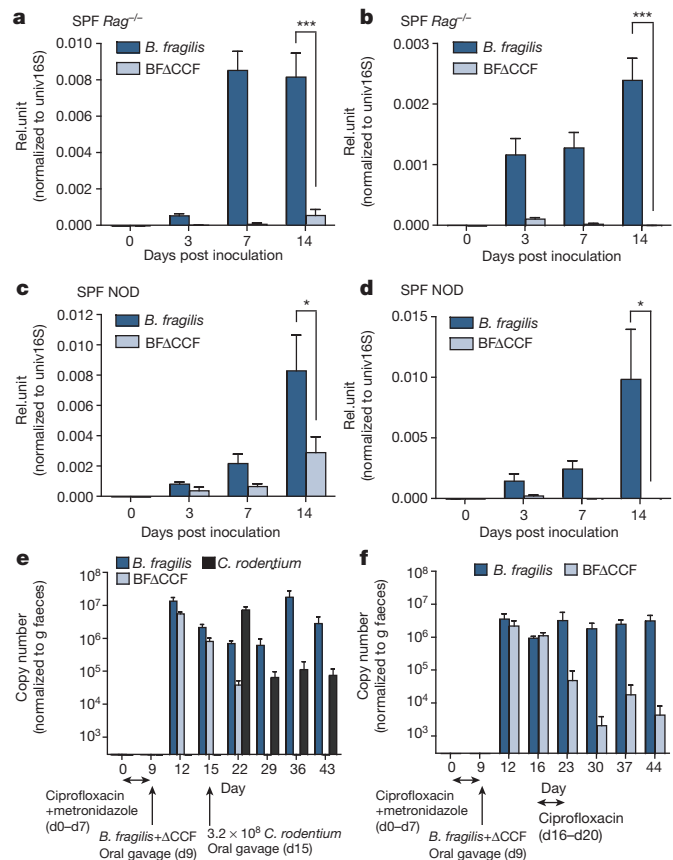
Building on the previous discovery that a population of *B. fragilis* associates with mucosal tissues<sup>22</sup>, we show that *ccfB–E* are preferentially expressed by bacteria in contact with the colon, with lower levels in caecal content and faeces (Fig. 2a). There is virtually no expression in laboratory culture. Thus, *in vivo* expression of *ccf* in gut tissue may be critical for colonization. Indeed, in contrast to laboratory-grown bacteria (see Fig. 1c), sustained colonization is conferred to bacteria recovered directly from animals (Supplementary Fig. 5). To examine regulation of gene expression, we deleted the  $\sigma$  factor *ccfA*, which led to highly reduced expression of all five genes during animal colonization (Supplementary Fig. 6a). Accordingly, germ-free mice mono-colonized with the *B. fragilis*  $\Delta$ ccfA mutant are permissive of super-colonization by wild-type *B. fragilis*, demonstrating a functional defect in the saturable niche occupancy (Supplementary Fig. 6b). On the basis of this tissue-associated expression pattern, we tested whether the CCF system promotes bacterial localization to mucosal tissue. Mice were mono-associated with either the wild-type or the *ccf* deletion strain, and both groups were subsequently challenged with wild-type *B. fragilis*. Twenty-four hours after challenge, we observed the same numbers for challenge strains in faeces of both groups (Fig. 2b first and second bars, and Supplementary Fig. 7a, b). By contrast, *ccf*-mutant-associated animals show higher levels of challenge strain in colon tissue, suggesting a colonization defect by the mutant strain specifically at the mucosal surface (Fig. 2b third and fourth bars, and Supplementary Fig. 7a, b). Therefore, CCF-mediated colonization fitness seems to involve physical association with the gut.

Recent studies have revealed microbial communities that colonize intestinal crypts of conventional mice in the absence of disease<sup>23</sup>, and we have shown that *B. fragilis* occupies the colonic crypts of mono-associated mice<sup>22</sup>. Discovering a role for *ccf* genes near mucosal tissue led us to explore the intriguing hypothesis that the CCF system mediates crypt occupancy. We mono-colonized mice with *B. fragilis* and visualized bacterial localization in colon tissue by whole-mount confocal microscopy. Indeed, wild-type *B. fragilis* co-localize with crypts from the ascending colon, appearing to be located in the centre of crypt opening. Notably, BF $\Delta$ CCF mono-associated mice display virtually no crypt occupancy (Fig. 2c and Supplementary Fig. 8). Colon cross-section imaging also reveals that only wild-type bacteria are crypt associated (Supplementary Fig. 9). Two-photon imaging of colon explants clearly demonstrates the presence of wild-type *B. fragilis* on the surface of the epithelium and inside the crypt. Although both wild-type and mutant strains of *B. fragilis* associate with the surface of the epithelium, only wild-type bacteria are able to penetrate deep into the colonic crypts of mice (Fig. 2d and Supplementary Video 1). Measuring the distance from the surface of the epithelium to bacterial signals in a survey of crypts reveals significantly greater tissue penetration by wild-type bacteria (Fig. 2e). Collectively, these data reveal that the CCF system allows *B. fragilis* to reside in a specific niche within crypts during steady-state colonization.

We next investigated the effects of the CCF system in the context of a complex microbiota. Wild-type *Bacteroides* species do not readily colonize most strains of specific pathogen-free (SPF) mice, namely BALB/c, Swiss Webster and C57BL/6, despite oral administration of high inocula (Supplementary Fig. 10a, c and data not shown). Furthermore, transfer of an SPF microbiota to mono-colonized mice leads to clearance of wild-type *B. fragilis* (Supplementary Fig. 10b, d). To overcome this obstacle, we tested various additional genetic backgrounds and empirically determined that C57BL/6 *Rag*<sup>−/−</sup> mice (which lack an adaptive immune system) and non-obese diabetic (NOD) mice can be stably colonized by *B. fragilis* with a single oral gavage. We introduced either wild-type or *ccf* mutant *B. fragilis* at equal inocula into separate groups of animals, and measured colonization. Only wild-type *B. fragilis* stably colonizes SPF *Rag*<sup>−/−</sup> mice, whereas BF $\Delta$ CCF establishes a significantly lower colonization in the gut (Fig. 3a). Co-inoculation of equal

numbers of wild-type and *ccf* mutant bacteria into *Rag*<sup>−/−</sup> mice also results in rapid clearance of the mutant strain from the gut (Fig. 3b), demonstrating a cell-intrinsic defect that could not be complemented *trans* by wild-type bacteria. NOD animals are also preferentially colonized by wild-type *B. fragilis* compared to *ccf* mutants in separate groups of animals (Fig. 3c) or in equal co-inoculation (Fig. 3d). These data show that deletion of the *ccf* genes compromises *B. fragilis* colonization of hosts with a complex microbiota.

During symbiosis with mammals, the microbiota may be confronted by rapid environmental changes with potentially adverse consequences to bacteria, such as enteric infections or antibiotic exposure. Gastroenteritis is commonly experienced by humans and is known to perturb the microbiota. To test whether resilience of *B. fragilis* colonization is CCF-dependent, we used *C. rodentium* infection of mice to mimic human gastrointestinal tract infection<sup>24</sup>. Using an antibiotic treatment protocol that does not sterilize the gut but promotes colonization of SPF mice by *Bacteroides*<sup>25</sup>, we were able to simultaneously colonize mice with equivalent levels of wild-type and *ccf* mutant bacteria. Mice were subsequently infected orally with *C. rodentium*, and colonization of *B. fragilis* was monitored. Wild-type bacteria decline in number at first, but return to maximal levels 3–4 weeks after infection (Fig. 3e). Importantly, the BF $\Delta$ CCF strain is completely cleared from the mouse



**Figure 3 | *B. fragilis* requires the *ccf* genes for stable and resilient colonization of mice.** **a**, Groups of SPF *Rag*<sup>−/−</sup> mice were gavaged with either wild-type *B. fragilis* or BF $\Delta$ CCF. **b**, SPF *Rag*<sup>−/−</sup> mice were given a 1:1 co-inoculum of wild-type *B. fragilis* and BF $\Delta$ CCF by single gavage. **c**, SPF NOD mice were gavaged with either wild-type *B. fragilis* or BF $\Delta$ CCF. **d**, SPF NOD mice were given a 1:1 co-inoculum of wild-type *B. fragilis* and BF $\Delta$ CCF by single gavage. **e**, SPF mice were co-associated with wild-type *B. fragilis* and BF $\Delta$ CCF, and infected with *C. rodentium*. **f**, SPF mice were co-associated with wild-type *B. fragilis* and BF $\Delta$ CCF, and given ciprofloxacin in drinking water for the time period shown. For all analyses, bacterial colonization levels were assessed by qRT-PCR from stool DNA ( $n = 4$  animals per group). Results are representative of at least two independent trials per experiment. Error bars indicate s.e.m. \* $P < 0.05$ ; \*\*\* $P < 0.001$ .

gut after gastroenteritis (Fig. 3e), but not when animals are left uninfected (Supplementary Fig. 11a). Next, we challenged mice that were co-colonized with wild-type and BFΔCCF with oral antibiotics and observe a colonization defect only in *ccf* mutant bacteria (Fig. 3f and Supplementary Fig. 11b). These results reveal that the CCF system establishes resilient colonization by gut *Bacteroides* following disruption of the microbiome. Finally, when SPF mice colonized with wild-type *B. fragilis* were given an antibiotic treatment that cleared faecal bacteria, crypt-associated microbial populations persisted (Supplementary Fig. 12), suggesting that symbiotic bacteria occupy a protected niche that creates a reservoir for stable gut colonization.

Co-evolution has bound microbes and man in an inextricable partnership, resulting in remarkable specificity and stability of the human microbiome<sup>2,3</sup>. Our findings reveal a novel pathway required for persistent colonization of the mammalian gut by the *Bacteroides*. Homology to the Sus family of proteins suggests a role for CCF in uptake and use of glycans. Although certain Sus-containing PULs in *B. thetaiotaomicron* mediate foraging of host mucus<sup>15</sup>, their contributions to microbial colonization have been previously described only during nutrient deprivation conditions<sup>20</sup>. Our discovery of CCF-dependent colonization in mice fed a nutrient-rich diet suggests a new role whereby *Bacteroides* evolved specific Sus-like systems to use non-dietary glycans during homeostasis. On the basis of findings that *ccf* genes are preferentially expressed in proximity to mucosal tissues and *B. fragilis* associates with colonic crypts, we find it likely that host factors may promote expression of the CCF system. In support of this notion, *N*-acetyl-D-lactosamine (LacNAc)—a component of host mucus—induces the *ccf* genes and its homologues in *B. fragilis* and *B. thetaiotaomicron*<sup>20</sup> (Supplementary Fig. 13). But because a closely related PUL (BF0227–31) responds to LacNAc but does not mediate saturable niche colonization (Supplementary Fig. 4), LacNAc alone may be an inducer but is not the substrate used by CCF systems. We propose that specific glycan structures within colonic crypts serve as nutrient sources for individual *Bacteroides* species, and that CCF systems provide a molecular mechanism for a hypothesis proposed decades ago, “that populations of most indigenous intestinal bacteria are controlled by substrate competition, i.e., that each species is more efficient than the rest in utilizing one or a few particular substrates and that the population level of that species is controlled by the concentration of these few limiting substrates”<sup>26</sup>. Future work will aim to identify the precise glycan(s) for CCF systems from various *Bacteroides*. Finally, our data suggest that the *ccf* genes encode a specific subset of PULs that evolved the novel activity of promoting stable and resilient colonization, and crypt-associated bacterial reservoirs may represent ‘founder’ cells that repopulate the gut following disruption of the microbiome by enteric infections or antibiotic exposure. Discovery of a molecular mechanism for colonization fitness by gut bacteria provides a glimpse into the evolutionary forces that have shaped the assembly and dynamics of the human microbiome.

## METHODS SUMMARY

All germ-free mice were bred and housed in flexible film isolators until 8 weeks of age, then transferred to microisolator cages and maintained with autoclaved food, bedding and water supplemented with gentamicin and erythromycin. Mice were mono-associated with gentamicin- and erythromycin-resistant *Bacteroides* strains by single oral gavage. Colonization level was determined over time by stool serial dilution plating on selective agar media. For qRT-PCR, total RNA was extracted from laboratory bacterial culture, faecal and caecal content (ZR Soil/Faecal RNA MicroPrep, Zymo Research) and colon tissues (Trizol, Invitrogen) from mono-associated animals, converted to first strand complementary DNA and analysed by qPCR using Power SYBR Green PCR Master Mix (Applied Biosystems). For colon whole-mount imaging, tissues were collected from germ-free or single strain mono-associated animals, fixed with 4% paraformaldehyde and stained with a *B. fragilis* specific antibody, DAPI and phalloidin. The colon crypts were visualized by confocal microscopy and two-photon microscopy. SPF mice were colonized with *B. fragilis* and/or BFΔCCF by single oral gavage and the colonization level was determined over time by stool DNA extraction (ZR Faecal DNA MiniPrep, Zymo Research) and qPCR using strain specific primers.

**Full Methods** and any associated references are available in the online version of the paper.

Received 6 September 2012; accepted 11 July 2013.

Published online 18 August 2013.

- McFall-Ngai, M. *et al.* Animals in a bacterial world, a new imperative for the life sciences. *Proc. Natl Acad. Sci. USA* **110**, 3229–3236 (2013).
- The Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207–214 (2012).
- Yatsunenko, T. *et al.* Human gut microbiome viewed across age and geography. *Nature* **486**, 222–227 (2012).
- Qin, J. *et al.* A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**, 59–65 (2010).
- Palmer, C., Bik, E. M., DiGiulio, D. B., Relman, D. A. & Brown, P. O. Development of the human infant intestinal microbiota. *PLoS Biol.* **5**, e177 (2007).
- Dethlefsen, L., Huse, S., Sogin, M. L. & Relman, D. A. The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. *PLoS Biol.* **6**, e280 (2008).
- Eckburg, P. B. *et al.* Diversity of the human intestinal microbial flora. *Science* **308**, 1635–1638 (2005).
- Bry, L., Falk, P. G., Midtvedt, T. & Gordon, J. I. A model of host-microbial interactions in an open mammalian ecosystem. *Science* **273**, 1380–1383 (1996).
- Xu, J. *et al.* A genomic view of the human-*Bacteroides thetaiotaomicron* symbiosis. *Science* **299**, 2074–2076 (2003).
- Mazmanian, S. K., Liu, C. H., Tzianabos, A. O. & Kasper, D. L. An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system. *Cell* **122**, 107–118 (2005).
- Mazmanian, S. K., Round, J. L. & Kasper, D. L. A microbial symbiosis factor prevents intestinal inflammatory disease. *Nature* **453**, 620–625 (2008).
- Round, J. L. & Mazmanian, S. K. Inducible Foxp3<sup>+</sup> regulatory T-cell development by a commensal bacterium of the intestinal microbiota. *Proc. Natl Acad. Sci. USA* **107**, 12204–12209 (2010).
- Ochoa-Repáraz, J. *et al.* Central nervous system demyelinating disease protection by the human commensal *Bacteroides fragilis* depends on polysaccharide A expression. *J. Immunol.* **185**, 4101–4108 (2010).
- Ley, R. E., Peterson, D. A. & Gordon, J. I. Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell* **124**, 837–848 (2006).
- Koropatkin, N. M., Cameron, E. A. & Martens, E. C. How glycan metabolism shapes the human gut microbiota. *Nature Rev. Microbiol.* **10**, 323–335 (2012).
- Martens, E. C., Koropatkin, N. M., Smith, T. J. & Gordon, J. I. Complex glycan catabolism by the human gut microbiota: the Bacteroidetes Sus-like paradigm. *J. Biol. Chem.* **284**, 24673–24677 (2009).
- Shipman, J. A., Berleman, J. E. & Salyers, A. A. Characterization of four outer membrane proteins involved in binding starch to the cell surface of *Bacteroides thetaiotaomicron*. *J. Bacteriol.* **182**, 5365–5372 (2000).
- Schauer, K., Rodionov, D. A. & de Reuse, H. New substrates for TonB-dependent transport: do we only see the ‘tip of the iceberg’? *Trends Biochem. Sci.* **33**, 330–338 (2008).
- Kawada, M. *et al.* Chitinase 3-like-1 enhances bacterial adhesion to colonic epithelial cells through the interaction with bacterial chitin-binding protein. *Lab. Invest.* **88**, 883–895 (2008).
- Martens, E. C., Chiang, H. C. & Gordon, J. I. Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. *Cell Host Microbe* **4**, 447–457 (2008).
- Sonnenburg, J. L. *et al.* Glycan foraging *in vivo* by an intestine-adapted bacterial symbiont. *Science* **307**, 1955–1959 (2005).
- Round, J. L. *et al.* The Toll-like receptor 2 pathway establishes colonization by a commensal of the human microbiota. *Science* **332**, 974–977 (2011).
- Pédon, T. *et al.* A crypt-specific core microbiota resides in the mouse colon. *MBio* **3**, e00116–12 (2012).
- Mundy, R., MacDonald, T. T., Dougan, G., Frankel, G. & Wiles, S. *Citrobacter rodentium* of mice and man. *Cell. Microbiol.* **7**, 1697–1706 (2005).
- Bloom, S. M. *et al.* Commensal *Bacteroides* species induce colitis in host-genotype-specific fashion in a mouse model of inflammatory bowel disease. *Cell Host Microbe* **9**, 390–403 (2011).
- Freter, R., Brickner, H., Botney, M., Clevon, D. & Aranki, A. Mechanisms that control bacterial populations in continuous-flow culture models of mouse large intestinal flora. *Infect. Immun.* **39**, 676–685 (1983).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank T. Thron and S. McBride for the maintenance of germ-free animals, J. Selicha for assisting with the experimental procedures and G. Chodaczek for help with confocal and two-photon microscopy. We are grateful to E. C. Martens and members of the Mazmanian laboratory for critical review of the manuscript. S.M.L. and G.P.D. were supported by a pre-doctoral training grant (GM007616). This work was supported by grants from the National Institutes of Health (GM099535 and DK078938) and the Crohn's and Colitis Foundation of America to S.K.M.

**Author Contributions** S.M.L. and S.K.M. conceived the project. S.M.L. performed most of the experiments; G.P.D., Z.M. and S.B. contributed data. S.M.L., G.P.D., Z.M., K.L. and S.K.M. interpreted the data. K.L. and S.K.M. secured funding. S.M.L. and S.K.M. wrote the manuscript. G.P.D., Z.M. and K.L. edited the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.K.M. ([sarkis@caltech.edu](mailto:sarkis@caltech.edu)).



## METHODS

**Bacterial strains, plasmids and culture conditions.** Bacterial strains and plasmids are described in Supplementary Table 1. *Bacteroides* strains were grown anaerobically at 37 °C for 2 days in brain heart infusion broth supplemented with 5 µg ml<sup>-1</sup> hemin and 0.5 µg ml<sup>-1</sup> vitamin K (BHIS), with gentamicin (200 µg ml<sup>-1</sup>), erythromycin (5 µg ml<sup>-1</sup>), chloramphenicol (10 µg ml<sup>-1</sup>) and tetracycline (2 µg ml<sup>-1</sup>) added where appropriate. *E. coli* JM109 containing recombinant plasmids were grown in luria broth (LB) with ampicillin (100 µg ml<sup>-1</sup>) or kanamycin (30 µg ml<sup>-1</sup>). *C. rodentium* DBS100 strain was grown in LB at 37 °C for 24 h. For the induction of *susC/D* homologues, *B. fragilis* and *B. thetaiotaomicron* were grown in minimal medium with either glucose or N-acetylglucosamine as the sole carbon source as described previously<sup>20</sup>.

**Mice.** 8–10-week old male and female germ-free Swiss Webster mice were purchased from Taconic Farms and bred in flexible film isolators. For gnotobiotic colonization experiments, germ-free mice were transferred to freshly autoclaved microisolator cages, fed *ad libitum* with a standard autoclaved chow diet and given autoclaved water supplemented with 10 µg ml<sup>-1</sup> of erythromycin and 100 µg ml<sup>-1</sup> of gentamicin. Male SPF C57BL/6 mice and Swiss Webster mice were purchased from Taconic Farms. Male SPF NOD/ShiLtJ mice and *Rag*<sup>-/-</sup> C57BL/6 mice were purchased from the Jackson Laboratory. No randomization or blinding was used to allocate experimental groups. Sample size and standard deviation were based on empirical data from pilot experiments. All procedures were performed in accordance with the approved protocols using IACUC guidelines of the California Institute of Technology.

**Construction of chromosomal library and screen.** Genomic DNA was isolated from overnight culture of *B. fragilis* using a commercial kit (Wizard Genomic DNA Purification Kit, Promega). 20 µg of genomic DNA was incubated with 4 U of Sau3AI for 5, 10, 15 or 20 min at 37 °C in 50 µl volume and the partially digested genomic DNA was separated by electrophoresis on 0.7% agarose gel. 9–10-kb fragment DNA was excised and recovered from the agarose gel (Zymoclean Gel DNA Recovery Kit, Zymo Research). Insert DNA was ligated to BglII site of plasmid vector (pFD340-catBII, Supplementary Table 1), transformed into *E. coli* and amplified on LB-ampicillin plate. Individual clones from the plasmid library were mobilized from *E. coli* to *B. vulgatus* by conjugal helper plasmid RK231 generating a library of *B. vulgatus* hosting *B. fragilis* chromosomal DNA fragments consisting of approximately ~2,100 clones. To screen the library *in vivo*, pools of 96 clones (10<sup>6</sup> c.f.u. of each clone) were gavaged into 22 germ-free Swiss Webster mice (10<sup>8</sup> c.f.u. per animal) pre-colonized with *B. vulgatus* pFD340 for 1–2 weeks. Two weeks after gavage, fresh faecal samples were plated on BHIS agar plate containing chloramphenicol to select for clones with colonization phenotype.

**Generation of *ccfA*, *ccfC*, *ccfD*, *ccfE*, *ccfC-E* (ΔCCF) and ΔBF0227–0229 deletion mutants.** ~2-kb DNA segments flanking the region to be deleted were PCR amplified using primers listed in Supplementary Table 2. Reverse primer of the left-flanking DNA and forward primer of the right-flanking DNA were designed to be partially complementary at their 5' ends by 18–21 bp. Fusion PCR was performed using the left and right flanking DNA (~300 ng each after gel purification) as DNA template and forward primer of the left-flanking DNA and reverse primer of the right-flanking DNA<sup>27</sup>. The fused PCR product was cloned into BamHI or SalI site of the *Bacteroides* conjugal suicide vector pNJR6 and mobilized into *B. fragilis*. Colonies selected for erythromycin resistance (Em<sup>r</sup>), indicating integration of the suicide vector into the host chromosome were passaged for 5 days and then plated on nonselective medium (BHIS). The resulting colonies were replica plated to BHIS containing Em, and Em<sup>s</sup> (erythromycin sensitive) colonies were screened by PCR to distinguish wild-type revertants from strains with the desired mutation. The same strategy was used to generate Δ*ccfC-E* deletion mutants from *B. vulgatus*.

**qRT-PCR.** Total RNA was extracted from mid-log phase bacterial culture using ZR Fungal/Bacterial RNA MiniPrep (Zymo Research), faeces and caecal content from mice using ZR Soil/Fecal RNA MicroPrep (Zymo Research), and mouse colon tissues after removing luminal content by gently scraping the mucosal surface and PBS rinse using Trizol (Invitrogen). Complementary DNA was made using an iSCRIPT cDNA synthesis kit per manufacturer's instructions (Bio-Rad). All qRT-PCR reactions were performed in ABI PRISM 7900HT Fast Real-Time PCR System (Applied Biosystems) using Power SYBR Green PCR Master Mix (Applied Biosystems). Gene-specific primers are described in Supplementary Table 2.

**Immunofluorescent staining of colon whole-mounts and frozen sections.** For whole-mount staining, colons were fixed in buffered 4% paraformaldehyde, washed with PBS and subjected to indirect immunofluorescence. Tissues were made permeable by incubation with 0.5% (wt/vol) saponin, 2% (vol/vol) FBS and 0.09% (wt/vol) azide in PBS for at least 18 h. The same buffer was used for subsequent incubations with antibodies. Colon fragments were incubated with a primary polyclonal chicken IgY anti-*B. fragilis* antibodies for 12–16 h at room

temperature followed by 1–2 h incubation at 37 °C. Following PBS washes, samples were reacted with goat anti-chicken IgY secondary antibodies (Alexa Fluor 488 or Alexa Fluor 633, 2 µg ml<sup>-1</sup>, Molecular Probes), fluorescently labelled phalloidin (fluorescein or AF568, 2 U ml<sup>-1</sup>, Molecular Probes) and DAPI (2 µg ml<sup>-1</sup>, Molecular Probes) for 1 h at room temperature. Tissues were mounted in Prolong Gold (Invitrogen) and allowed to cure for at least 48 h before imaging. In some experiments, anti-*B. fragilis* antibodies were pre-absorbed on tissue fragments derived from either germ-free mice (up to 18 h at room temperature) or SPF mice (1 h at room temperature).

For frozen sections, colon tissues were embedded in OCT Compound (Sakura Finetek), frozen on dry ice and stored at -80 °C. Frozen blocks were cut with a thickness of 10 µm using a Microm HM505E cryostat, and sections were collected on positively charged slides (Fisher Scientific) for staining. Slides were fixed with 4% buffered paraformaldehyde for 10 min and washed 2 × 10 min with PBS. Tissue sections were blocked with 10% normal goat serum and 0.5% bovine serum albumin in PBS for 1 h at room temperature. Sections were incubated with anti-*B. fragilis* antibodies for at least 8 h at 4 °C, washed twice for 10 min with PBS, reacted with secondary reagents and mounted as described above. In some experiments, anti-*B. fragilis* antibodies were pre-absorbed for 1 h at room temperature on tissue sections derived from germ-free mice.

**Fluorescence microscopy.** An SP5 resonant laser-scanning confocal and two-photon microscope (both scanning heads mounted on the same DM 6000 upright microscope, Leica Microsystems) with a 40× oil objective (numerical aperture 1.4) or 63× oil objective (numerical aperture 1.4) were used for fluorescence microscopy. Images used for three-dimensional reconstructions were acquired using dual confocal-two-photon mode. For confocal imaging, 488-nm and 543-nm excitation wavelengths were used for Alexa Fluor 488-labelled bacteria and Alexa Fluor 568-labelled phalloidin, and signals were detected with internal photomultiplier tubes. Two-photon imaging was performed with four nondescanned detectors (Leica Microsystems) and a Chameleon Ultra Ti: Sapphire laser (Coherent) tuned at 700–800 nm for acquisition. Emitted fluorescence was split with three dichroic mirrors (496 nm, 560 nm and 593 nm) and passed through an emission filter (Semrock) at 585/40 nm. Images (512 × 512) acquired with a 0.5-µm Z step were smoothed by median filtering at kernel size 3 × 3 pixels. Three-dimensional reconstructions of crypts and bacteria were performed using Imaris software (version 7.5.1 × 64; Bitplane AG). Crypt structures were visualized by DAPI and phalloidin signals. Images used for quantification were acquired with FluoView FV10i confocal microscope (Olympus) using 60× (numerical aperture 1.35) oil objective.

**Image analysis.** For bacterial localization with respect to the epithelial layer, frames of 512 × 512 pixels were acquired with 1-µm Z steps in the crypt length axis. Images were processed using ImageJ software (NIH). Background was subtracted (rolling ball method), images were smoothed by median filtering (3 × 3 pixels), segmented by threshold and position of the signal in the Z stack was recorded. Data did not follow normal distribution and were analysed by non-parametric two-sided Mann–Whitney U-test.

For quantification of crypt-associated bacterial signals from antibiotic-treated animals, stacks of 512 × 512 pixels by eight frames (1 µm per frame) were flattened by maximum intensity projection and filtered by median (3 × 3 kernel size). Images were segmented by thresholding. Number of positive spots per 1,000 µm<sup>2</sup> and area occupied by individual spots were analysed. Data were not normally distributed and were analysed by Mann–Whitney or Kruskal–Wallis followed by Dunn's multiple comparisons test where appropriate. 11–13 stacks/group were examined. Total area that was analysed within the group of stacks was between 0.08–0.2 mm<sup>2</sup>.

**Gnotobiotic animal colonization experiments.** 8–12-week-old germ-free Swiss Webster mice were gavaged once with a 100 µl of bacterial suspension for mono-association (~10<sup>8</sup> c.f.u. of each bacterial strain collected from a log-phase culture and re-suspended in PBS with 1.5% NaHCO<sub>3</sub>). For sequential colonization, germ-free mice were mono-associated with an initial strain for 6–7 days and subsequently gavaged with a 100 µl suspension of a challenge strain. All *Bacteroides* strains used to colonize germ-free animals were resistant to gentamicin inherently, and to erythromycin by plasmid. Unless otherwise indicated, the initial strains carried pFD340-*cat* (chloramphenicol resistant; Cm<sup>r</sup>) and the challenge strains, pFD340-*tetQ* (tetracycline resistant; Tet<sup>r</sup>). For horizontal transfer by encounter experiment, two single-housed mice that were mono-associated with either wild-type *B. fragilis* pFD340-*tetQ* or BFΔCCF pFD340-*cat* for at least 3 weeks were co-housed in a fresh sterile cage for 4 h and then separated. At each time point, fresh faecal samples were collected, weighed, homogenized and serially diluted in PBS (or BHI broth) for plating on selective media to determine bacterial c.f.u. per g of faeces.

**SPF animal colonization experiments.** 7–8-week-old male SPF mice (C57BL/6, Swiss Webster, NOD and *Rag*<sup>-/-</sup>) were given a single inoculum of 1 × 10<sup>8</sup> c.f.u. of either wild-type *B. fragilis*, BFΔCCF or 1:1 mixture of the two strains by oral



gavage. At each time point, bacterial genomic DNA from faecal samples were isolated using a commercial kit (ZR Fecal DNA MiniPrep, Zymo Research) following the manufacturer's instructions and the relative densities of bacteria were determined by qPCR using strain-specific primers (Supplementary Table 2).

**C. rodentium infection.** 8-week-old female SPF Swiss Webster mice were treated with metronidazole ( $100 \text{ mg kg}^{-1}$ ) by oral gavage every 24 h and ciprofloxacin dissolved in drinking water ( $0.625 \text{ mg ml}^{-1}$ ; Hikma Pharmaceuticals) for 7 days; mice were transferred to a fresh sterile cage every 2 days. 2 days after the cessation of antibiotic treatment, mice were orally gavaged with a single inoculum of 1:1 mixture of wild-type *B. fragilis* and BFΔCCF ( $\sim 5 \times 10^8$  c.f.u. total per animal). 6–7 days after *B. fragilis* gavage, mice were either infected orally with  $\sim 5 \times 10^8$  c.f.u. of overnight culture *C. rodentium* or PBS-gavaged as control. The relative densities of bacteria were determined by faecal bacterial DNA extraction and qPCR.

**Antibiotic treatment.** 8-week-old female SPF Swiss Webster mice were treated with metronidazole ( $100 \text{ mg kg}^{-1}$ ) by oral gavage every 24 h and ciprofloxacin dissolved in drinking water ( $0.625 \text{ mg ml}^{-1}$ ) for 7 days; mice were transferred to a fresh sterile cage every 2 days. 2 days after the cessation of antibiotic treatment, mice were orally gavaged with a single inoculum of 1:1 mixture of wild-type *B. fragilis* and BFΔCCF ( $\sim 5 \times 10^8$  c.f.u. total per animal). 6–7 days after *B. fragilis*

gavage, one group of mice were treated with ciprofloxacin for 4 days dissolved in drinking water ( $1 \text{ mg ml}^{-1}$ ) and another group were left untreated. The relative densities of bacteria were determined by faecal bacterial DNA extraction and qPCR.

**Antibiotic treatment for colon whole-mount imaging.** 8-week-old female SPF Swiss Webster mice were treated with metronidazole ( $100 \text{ mg kg}^{-1}$ ) by oral gavage every 24 h and ciprofloxacin dissolved in drinking water ( $0.625 \text{ mg ml}^{-1}$ ) for 7 days; mice were transferred to a fresh sterile cage every 2 days. 2 days after the cessation of antibiotic treatment, mice were orally gavaged with a  $100 \mu\text{l}$  inoculum of *B. fragilis* ( $\sim 10^8$  c.f.u.) or PBS. 7 days after bacterial gavage (day 16), PBS or *B. fragilis* inoculated mice were treated with ciprofloxacin in drinking water ( $1 \text{ mg ml}^{-1}$ ) for 7 days and one group of *B. fragilis* inoculated mice were left untreated. At the end of the ciprofloxacin treatment (day 23), faeces were collected for stool DNA extraction and colon tissues were collected and fixed with 4% paraformaldehyde for whole-mount imaging.

27. Wurch, T., Lestienne, F. & Pauwels, P. J. A modified overlap extension PCR method to create chimeric genes in the absence of restriction enzymes. *Biotechnol. Tech.* **12**, 653–657 (1998).

# Mechanism of *Trypanosoma brucei gambiense* resistance to human serum

Pierrick Uzureau<sup>1</sup>, Sophie Uzureau<sup>1</sup>, Laurence Lecordier<sup>1</sup>, Frédéric Fontaine<sup>1</sup>, Patricia Tebabi<sup>1</sup>, Fabrice Homblé<sup>2</sup>, Axelle Grélard<sup>3</sup>, Vanessa Zhendre<sup>3</sup>, Derek P. Nolan<sup>4</sup>, Laurence Lins<sup>5</sup>, Jean-Marc Crowet<sup>5</sup>, Annette Pays<sup>1</sup>, Cécile Felu<sup>1</sup>, Philippe Poelvoorde<sup>1</sup>, Benoit Vanhollebeke<sup>1</sup>, Søren K. Moestrup<sup>6</sup>, Jeppe Lyngsø<sup>7</sup>, Jan Skov Pedersen<sup>7</sup>, Jeremy C. Mottram<sup>8</sup>, Erick J. Dufourc<sup>3</sup>, David Pérez-Morga<sup>9</sup> & Etienne Pays<sup>10</sup>

The African parasite *Trypanosoma brucei gambiense* accounts for 97% of human sleeping sickness cases<sup>1</sup>. *T. b. gambiense* resists the specific human innate immunity acting against several other tsetse-fly-transmitted trypanosome species such as *T. b. brucei*, the causative agent of nagana disease in cattle. Human immunity to some African trypanosomes is due to two serum complexes designated trypanolytic factors (TLF-1 and -2), which both contain haptoglobin-related protein (HPR) and apolipoprotein LI (APOL1)<sup>2–4</sup>. Whereas HPR association with haemoglobin (Hb) allows TLF-1 binding and uptake via the trypanosome receptor TbHpHbR (ref. 5), TLF-2 enters trypanosomes independently of TbHpHbR (refs 4, 5). APOL1 kills trypanosomes after insertion into endosomal/lysosomal membranes<sup>2,6,7</sup>. Here we report that *T. b. gambiense* resists TLFs via a hydrophobic  $\beta$ -sheet of the *T. b. gambiense*-specific glycoprotein (TgsGP)<sup>8</sup>, which prevents APOL1 toxicity and induces stiffening of membranes upon interaction with lipids. Two additional features contribute to resistance to TLFs: reduction of sensitivity to APOL1 requiring cysteine protease activity, and TbHpHbR inactivation due to a L210S substitution. According to such a multifactorial defence mechanism, transgenic expression of *T. b. brucei* TbHpHbR in *T. b. gambiense* did not cause parasite lysis in normal human serum. However, these transgenic parasites were killed in hypohaptoglobinaemic serum, after high TLF-1 uptake in the absence of haptoglobin (Hp) that competes for Hb and receptor binding. TbHpHbR inactivation preventing high APOL1 loading in hypohaptoglobinaemic serum may have evolved because of the overlapping endemic area of *T. b. gambiense* infection and malaria, the main cause of haemolysis-induced hypohaptoglobinaemia in western and central Africa<sup>9</sup>.

The protozoan flagellate *Trypanosoma brucei brucei* infects many mammals, except some primates including humans, where it is lysed by the trypanolytic protein APOL1 (refs 2, 6, 7). APOL1 is associated with two different serum complexes termed TLF-1 and TLF-2 (refs 2–4). TLF-1 is a subset of high-density lipoprotein particles containing HPR, which allows these particles to enter trypanosomes through the Hp–Hb receptor TbHpHbR<sup>5</sup>. TLF-2 is an IgM-rich complex that also contains HPR (ref. 3). Whereas the Hp–Hb complex competitively inhibits TLF-1 binding to TbHpHbR, it does not interfere with TLF-2 uptake<sup>4,10</sup>. Two *T. brucei* subspecies, termed *T. b. rhodesiense* and *T. b. gambiense*, can resist APOL1 and therefore infect humans, in eastern and western and central Africa, respectively<sup>1,11</sup>.

Neutralization of APOL1 by *T. b. rhodesiense* results from interaction of the APOL1 carboxy-terminal helix with a specific shortened variant surface glycoprotein (VSG) termed serum resistance-associated protein

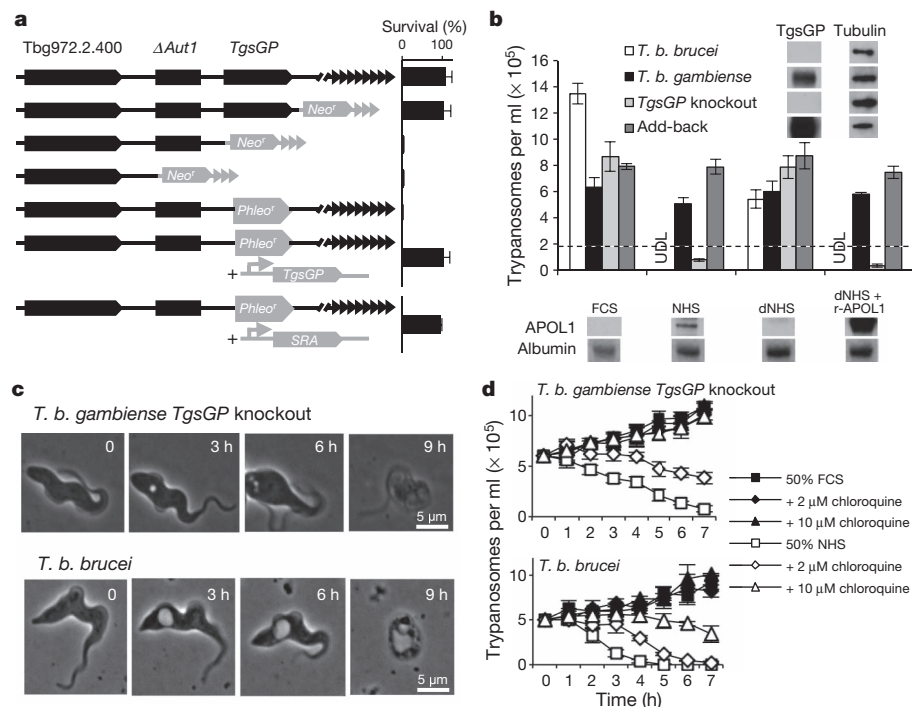
(SRA)<sup>6,12,13</sup>. Regarding *T. b. gambiense*, a search for shortened VSGs identified *T. gambiense*-specific glycoprotein (TgsGP)<sup>8</sup> (Supplementary Fig. 1). TgsGP is present in about  $2.2 \times 10^5$  copies per parasite or 50-fold less than the regular VSG (Supplementary Fig. 2). Like the VSG-like transferrin receptor<sup>14</sup>, TgsGP lacks the VSG C-terminal domain, contains a putative signal sequence for membrane anchoring by glycosylphosphatidylinositol (GPI) and localizes in the endocytic compartment<sup>8</sup> (Supplementary Fig. 3). Except for being VSG-like, TgsGP is unrelated to SRA (Supplementary Fig. 4), and in contrast to SRA, this gene cannot confer normal human serum (NHS) resistance to *T. b. brucei*<sup>8</sup> (Supplementary Fig. 5). TgsGP is absolutely specific to *T. b. gambiense*<sup>15–17</sup>, but TgsGP-related sequences and a TgsGP ancestor (Tb10.v4.0178) are present in *T. b. brucei*<sup>16,17</sup> (Supplementary Figs 4 and 6). TgsGP is located in the sub-telomeric region of chromosome 2, following DNA rearrangement that interrupted an *Aut1* allele<sup>16</sup>.

Progressive deletion of the TgsGP sub-telomeric region revealed that TgsGP is necessary for resistance to NHS (Fig. 1a, b and Supplementary Fig. 7). Neither the region between TgsGP and the telomere nor the truncated *Aut1* were implicated in resistance (Fig. 1a). Transgenic re-expression of TgsGP in TgsGP knockout parasites restored NHS resistance (Fig. 1a, b). TgsGP deletion only mildly affected parasite growth *in vitro* or in mice (Supplementary Fig. 8). Either parasite transfection with SRA, or SRA-mediated serum depletion of APOL1, restored growth of TgsGP knockout parasites in NHS, whereas recombinant APOL1 (r-APOL1) restored TgsGP knockout trypanolysis in APOL1-depleted NHS (Fig. 1a, b). Therefore, TgsGP conferred resistance to APOL1. As was observed with *T. b. brucei*<sup>2,7</sup>, NHS-mediated lysis of TgsGP knockout parasites was linked to swelling of endosomal vesicles (Fig. 1c) and was inhibited by chloroquine, a lysosomotropic agent interfering with endosomal acidification (Fig. 1d). However, TgsGP knockout parasite lysis was characterized by smaller vesicular swelling and delayed kinetics, indicating reduced APOL1 activity in *T. b. gambiense* (Fig. 1c, d and Supplementary Fig. 9).

TgsGP and r-APOL1 were found trafficking together from the flagellar pocket through the endocytic pathway, before finally reaching vesicles also accumulating transferrin (Fig. 2a and Supplementary Fig. 3a). However, despite the co-localization of TgsGP and APOL1, we failed to detect any high-affinity interaction between the two proteins, either in cellular extracts or *in vitro*, and no alternative protein associated with TgsGP (Supplementary Table 1).

Deletion of the GPI signal sequence did not affect TgsGP subcellular targeting or the ability to confer resistance (Supplementary Fig. 10 and Fig. 2b, Del1). The C-terminal sequence and the central domain predicted to be surface exposed<sup>18</sup> (Supplementary Fig. 1a) were also

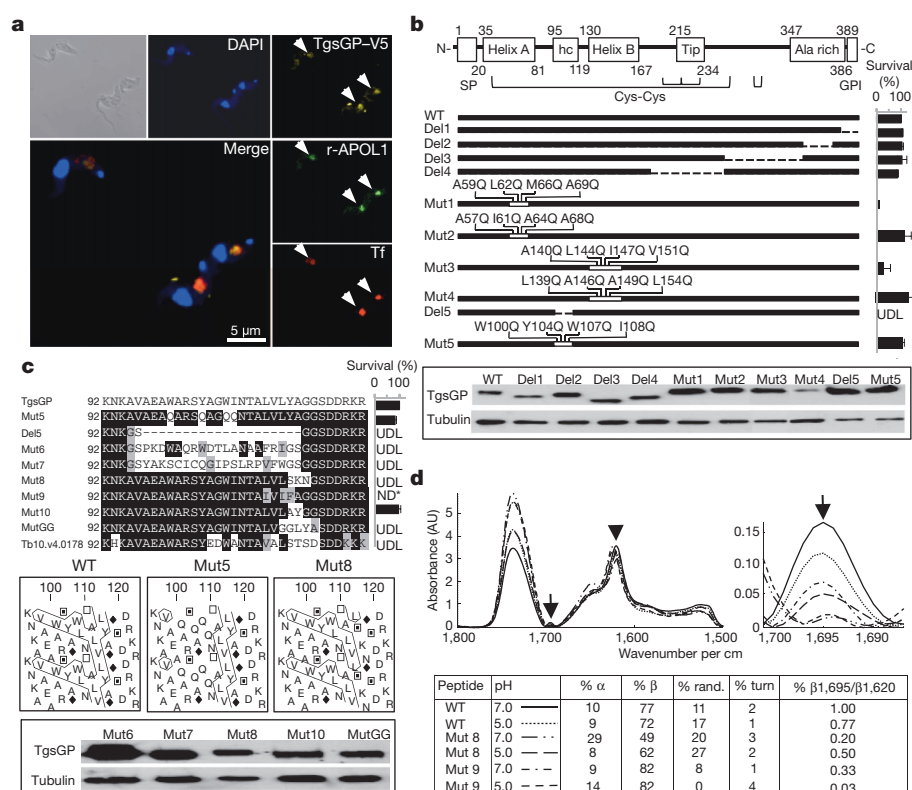
<sup>1</sup>Laboratory of Molecular Parasitology, IBMM, Université Libre de Bruxelles (ULB), 12 rue des Prof. Jeener et Brachet, B-6041 Gosselies, Belgium. <sup>2</sup>Structure and Function of Biological Membranes, Université Libre de Bruxelles, B-1050 Brussels, Belgium. <sup>3</sup>Institute of Chemistry & Biology of Membranes & Nanobiology, UMR 5248, CNRS, Université Bordeaux, Institut Polytechnique Bordeaux, F-33600 Pessac, France. <sup>4</sup>Molecular Parasitology Group, School of Biochemistry and Immunology, Trinity College Dublin, Dublin 2, Ireland. <sup>5</sup>Center of Numerical Molecular Biophysics, Université de Liège, B-5030 Gembloux, Belgium. <sup>6</sup>Department of Biomedicine, University of Aarhus, DK-8000 Aarhus, Denmark. <sup>7</sup>Interdisciplinary Nanoscience Center (iNANO) and Department of Chemistry, University of Aarhus, DK-8000 Aarhus, Denmark. <sup>8</sup>Wellcome Trust Centre for Molecular Parasitology, University of Glasgow, Glasgow G12 8TA, UK. <sup>9</sup>Center for Microscopy and Molecular Imaging (CMMI) and Laboratory of Molecular Parasitology, IBMM, Université Libre de Bruxelles, B-6041 Gosselies, Belgium. <sup>10</sup>Walloon Excellence in Life sciences and Biotechnology (WELBIO) and Laboratory of Molecular Parasitology, IBMM, Université Libre de Bruxelles, B-6041 Gosselies, Belgium.



**Figure 1 | TgsGP-mediated resistance to NHS.** **a**, *T. b. gambiense* resistance to NHS following deletions in the TgsGP-containing telomere of chromosome 2.  $\Delta$ Aut1, truncated Aut1 gene; *Neo*<sup>r</sup>/*Phleo*<sup>r</sup>, genes for neomycin or phleomycin resistance; triangles, telomeric repeats. Error bars represent standard deviation (s.d., 3 replicates;  $n = 3$ ). **b**, Trypanosome growth in 40% serum. dNHS, APOL1-depleted NHS; FCS, fetal calf serum; tubulin/albumin, control proteins; UDL, under detection limit. Error bars represent s.d. (3 replicates;  $n = 4$ ). **c**, Trypanosome lysis in 50% NHS (representative cell from a microscope field;  $n = 4$ ). **d**, Trypanosome growth with or without chloroquine. Error bars represent s.d. (3 replicates;  $n = 3$ ).

dispensable for TgsGP function (Fig. 2b, Del2–4). In contrast, mutations affecting hydrophobic interactions between heptad repeats of the amino-terminal amphipathic helices A and B<sup>19</sup> (Supplementary Fig. 1b, Mut1–4) inactivated TgsGP despite normal localization and global folding of the mutant proteins (Fig. 2b and Supplementary Figs 10 and 11). Therefore, the double-helical structure of the TgsGP N-terminal domain is important for activity. The linker between the two helices exhibits a type B VSG-specific length (Supplementary Fig. 1) and contains a hydrophobic sequence (residues 95–119) with  $\alpha$ -helical

and  $\beta$ -sheet components as indicated by both hydrophobic cluster analysis<sup>20</sup> (Fig. 2c) and infrared spectroscopy of the synthetic 92–124 peptide, which also revealed the  $\beta$ -sheet ability to form antiparallel structures (Fig. 2d). Sequence deletion or replacement with the corresponding stretch from TgsGP-related proteins or with irrelevant sequences led to loss of function (Fig. 2c, Del5, Mut6, Mut7). Whereas disruption of the hydrophobic character of the  $\alpha$ -helix did not prevent TgsGP activity (Fig. 2b, c, Mut5), reduction of the  $\beta$ -sheet hydrophobicity was sufficient to affect both TgsGP function and



**Figure 2 | TgsGP component responsible for resistance to NHS.** **a**, Immunofluorescence detection of *in situ*-V5-tagged TgsGP compared to r-APOL1 and transferrin (Tf) after 6 h incubation with cysteine protease inhibitor (representative cells from a microscope field;  $n = 4$ ). **b**, TgsGP knockout resistance to NHS after transfection with TgsGP constructs (boxed panel shows Western blot detection of TgsGP; tubulin is loading control). A scheme of TgsGP is shown above. Ala rich, alanine-rich region; GPI, GPI signal sequence; hc, hydrophobic cluster; SP, signal peptide; tip, surface-exposed loops. Error bars represent s.d. (3 replicates;  $n = 3$ ). **c**, Same legend as **b**, but focusing on the TgsGP inter-helical region (3 replicates;  $n = 3$ ). ND\*, not determined owing to growth inhibition. Plots of hydrophobic cluster analysis<sup>20</sup> are shown for some peptides. The boxed panel shows Western blot detection of TgsGP (tubulin is loading control). **d**, Infrared spectra of TgsGP peptides in soy lipids ( $n = 4$ ). The arrow and arrowhead, respectively, identify the peaks of antiparallel  $\beta$ -sheet and  $\beta$ -strand. AU, arbitrary units.



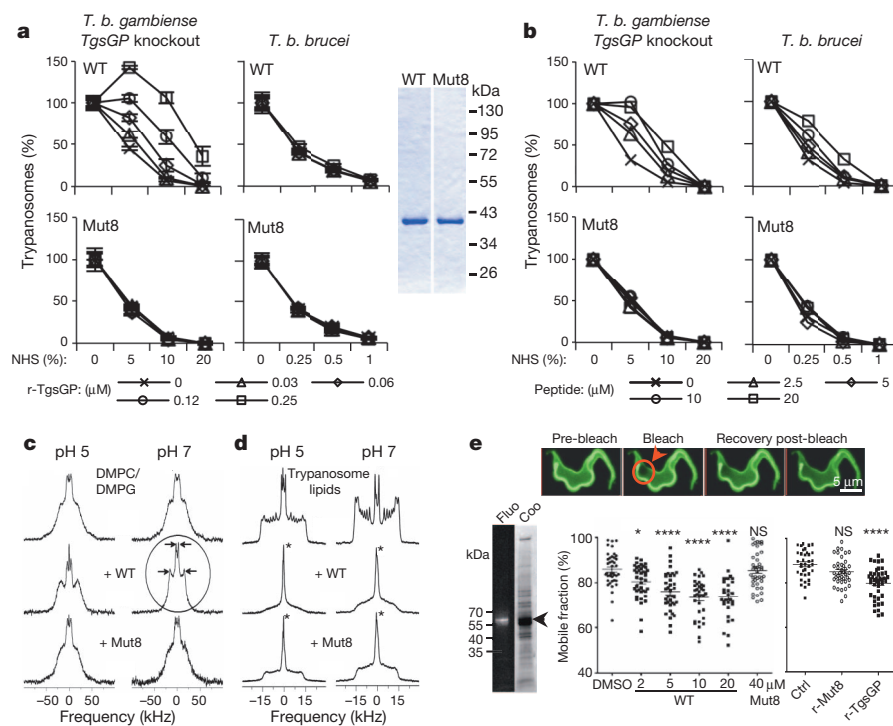
localization despite correct protein folding (Fig. 2c, d, Mut8, Tb10.v4.0178; Supplementary Figs 10 and 11). Alterations of the  $\beta$ -sheet sequence conserving its hydrophobicity but reducing the antiparallel potential (LVLY $\rightarrow$ IVIF, Mut9) led to growth inhibition *in vivo* (Fig. 2c, d). In contrast, minor changes such as YA $\rightarrow$ AY (Mut10) conserved full activity (Fig. 2c).

Incubation with recombinant TgsGP devoid of both N- and C-terminal signal peptides (r-TgsGP) provided dose-dependent protection of TgsGP knockout parasites against NHS, whereas r-Mut8 did not (Fig. 3a and Supplementary Fig. 12). Furthermore, the synthetic 92–124 peptide, but not its Mut8 version, conferred resistance to NHS in both TgsGP knockout parasites and *T. b. brucei* (Fig. 3b). With respect to r-TgsGP, the wild-type peptide required a 70-fold higher concentration to confer similar NHS protection (Supplementary Fig. 13a), as expected given the influence of the flanking  $\alpha$ -helices on TgsGP function (Fig. 2b, Mut1 and Mut3). Changes conserving hydrophobicity but not antiparallel organization largely inactivated both peptide and recombinant protein (Fig. 2d and Supplementary Fig. 13b, c, Mut9). Shifting two glycine residues inside the hydrophobic  $\beta$ -sheet, scrambling the sequence or deleting the terminal charged residues equally inactivated the peptide or recombinant protein (Fig. 2c; Supplementary Fig. 13b–d, MutGG, MutScr and MutDel). Thus, the hydrophobicity, antiparallel  $\beta$ -sheet organization and terminal basic amino acids of the inter-helical peptide are all important for resistance to APOL1.

In parasite cryosections, TgsGP was detected along membranes close to the flagellar pocket (Supplementary Fig. 3b). Assuming a size similar to the VSG (30 nm<sup>2</sup> per dimer), the  $2.2 \times 10^5$  TgsGP copies would cover 3  $\mu$ m<sup>2</sup> or about 7% of the endosomal surface<sup>21</sup>, in

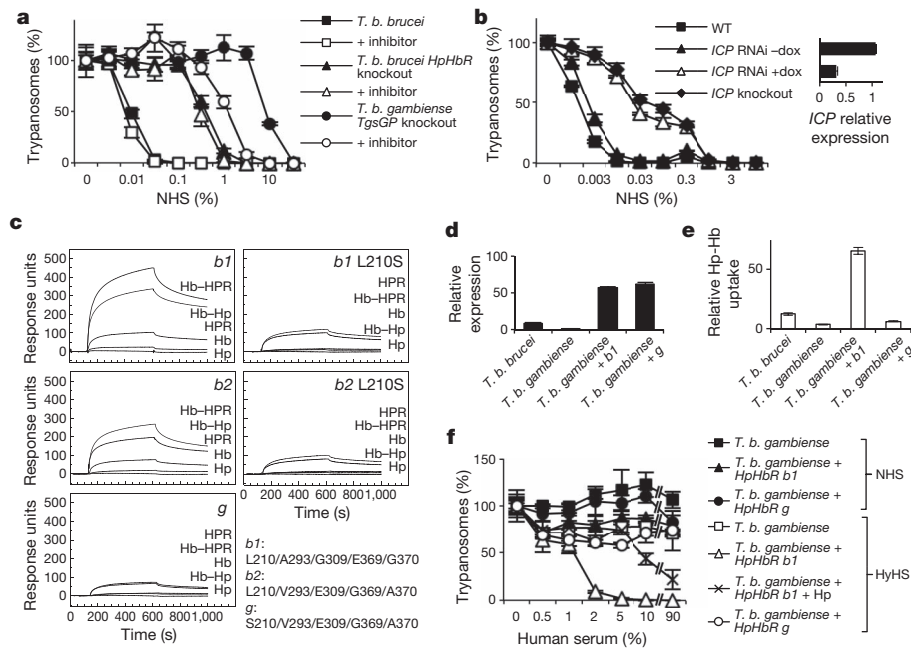
accordance with confocal microscopy measurements (Supplementary Fig. 3c). As determined by solid-state nuclear magnetic resonance (NMR), the wild-type peptide strongly interacted with lipid membranes and induced marked membrane stiffening, in contrast to Mut8, Mut9, MutGG and MutScr (Fig. 3c and Supplementary Figs 14–17). The wild-type effects involved electrostatic interactions between the positively charged peptide and lipid phosphate groups<sup>22</sup> (Supplementary Fig. 16). In liposomes made of trypanosome lipids, the wild-type and Mut8 peptides increased membrane curvature, with a stronger effect in the former case especially in acidic conditions (Fig. 3d). Therefore, TgsGP could confer resistance to APOL1 by stiffening and increased curvature of endosomal membranes. Moreover, measurements of fluorescence recovery after photobleaching (FRAP) indicated that r-TgsGP, but not r-Mut8, or the wild-type inter-helical TgsGP peptide, but not the Mut8, Mut9, MutGG and MutScr peptides, can decrease membrane fluidity as reflected by reduced lateral mobility of VSG on the surface of live trypanosomes (Fig. 3e and Supplementary Fig. 18).

NHS-mediated lysis was slower for TgsGP knockout parasites than for *T. b. brucei* (Fig. 1d). This was due to lower sensitivity to either NHS or r-APOL1, even compared to *TbHpHbR* knockout *T. b. brucei* (Fig. 4a and Supplementary Fig. 19). This differential NHS sensitivity was reversed by inhibiting cysteine protease activity (Fig. 4a). That cysteine proteases influence sensitivity to APOL1 was demonstrated by increased resistance of *T. b. brucei* to NHS after either deletion or RNA interference (RNAi)-mediated downregulation of the inhibitor of cysteine peptidases (*ICP*) gene (Fig. 4b). As compared to *T. b. brucei*, expression of cathepsins was not higher in *T. b. gambiense* (Supplementary Figs 20 and 21), but the pH of the endocytic



**Figure 3 | Activity of r-TgsGP and synthetic TgsGP peptides.** **a**, TgsGP knockout and *T. b. brucei* growth in NHS with or without r-TgsGP (right panel shows Coomassie-stained r-proteins; *n* = 2); 100% indicates trypanosome density without NHS. Error bars represent s.d. (3 replicates; *n* = 3). **b**, Same legend as **a**, but with or without TgsGP peptides (sequences in Fig. 2c). **c**, Solid-state NMR spectra of deuterated dimyristoylphosphatidylcholine (DMPC)/dimyristoylphosphatidylglycerol (DMPG) membranes interacting with TgsGP peptides at 20 °C (*n* = 3). Spectra narrowing and sharpening (arrows) demonstrate membrane interaction. **d**, Solid-state NMR spectra of deuterated palmitoyloleoylphosphatidylcholine (POPC) embedded in

trypanosome membrane lipids (*n* = 3). Spectral narrowing (asterisk) reflects increase in membrane curvature<sup>29</sup>. **e**, FRAP analysis of surface-labelled *T. b. brucei* with or without TgsGP peptides (in DMSO) or r-TgsGP. Top: the photobleached surface is encircled (image acquisition intervals: 500 ms). Left: SDS-PAGE of surface-labelled trypanosome extracts (fluor, fluorescence at 470 nm; Coomassie staining; arrowhead indicates VSG). Right: percentage of mobile VSG<sup>30</sup>; statistical significance was determined with one-way ANOVA with pairwise comparison Tukey's test (NS, not significant; single and quadruple asterisks refer to *P* values of 0.0473 and <0.0001, respectively). Error bars represent s.d. (40 replicates; *n* = 2).



**Figure 4 | Features linked to *T. b. gambiense* resistance to NHS.**

**a**, Trypanosome resistance to NHS with or without cysteine protease inhibitor. Error bars represent s.d. (3 replicates;  $n = 3$ ). **b**, Sensitivity of ICP knockout/ RNAi *T. b. brucei* to NHS. ICP expression was determined by qRT-PCR. Error bars represent s.d. (3 replicates;  $n = 3$ ). **c**, Ligand binding properties of TbHpHbR variants (*b1/b2*, *T. b. brucei* alleles; *g*, *T. b. gambiense* allele), measured by surface plasmon resonance. **d**, TbHpHbR expression after

transfection of TbHpHbR variants, determined by qRT-PCR. Error bars represent s.d. (3 replicates;  $n = 2$ ). **e**, Hp-Hb uptake after transfection of TbHpHbR variants, determined by flow cytometry with fluorescent Hp. Error bars represent s.d. (3 replicates;  $n = 2$ ). **f**, Survival of TbHpHbR-transfected *T. b. gambiense* in NHS or hypohaptoglobinaemic human serum (HyHS) with or without Hp. Error bars represent s.d. (3 replicates;  $n = 3$ ).

compartment was lower ( $4.85 \pm 0.05$  versus  $5.34 \pm 0.09$ ; Supplementary Table 2). Such pH difference can influence *T. brucei* cathepsin activity<sup>23</sup>. A lower pH in early endosomes and subsequent earlier endosomal protease activity could secure *T. b. gambiense* resistance to NHS by accelerating APOL1 degradation. In *T. b. brucei*, the absence of such an effect could explain the inefficacy of transfected TgsGP or r-TgsGP, even in TbHpHbR knockout parasites (Supplementary Fig. 5).

Another difference between *T. b. gambiense* and *T. b. brucei* concerns the TLF-1 receptor TbHpHbR (ref. 5). As assayed in *T. b. brucei*, the *T. b. gambiense* receptor was much less active<sup>24–26</sup>. Accordingly, whereas complementation of TbHpHbR knockout *T. b. brucei* with *T. b. brucei* TbHpHbR alleles (*b1* or *b2*) enhanced trypanolysis in NHS, transfection with *T. b. gambiense* TbHpHbR (*g*) did not (Supplementary Fig. 23). Receptor inactivation resulted from strong reduction of affinity for either Hp-Hb or HPR-Hb due to the L210S mutation (Fig. 4c), a mutation selectively conserved among *T. b. gambiense* isolates<sup>25–27</sup>. Complementation of *T. b. gambiense* with TbHpHbR *b1* restored Hp-Hb uptake (Fig. 4d, e); however, even provided with a functional TbHpHbR these parasites fully resisted NHS (Fig. 4f). In contrast, the same parasites were lysed in hypohaptoglobinaemic serum, in a process reversed by Hp addition (Fig. 4f). As hypohaptoglobinaemia strongly favours TLF-1 uptake over TLF-2 uptake<sup>4,5,10</sup>, it seems that TgsGP-mediated resistance to APOL1 can be bypassed by efficient TLF-1 uptake. Therefore, TbHpHbR down-regulation could have promoted *T. b. gambiense* survival in regions with high prevalence of hypohaptoglobinaemia. Such is the case in malaria endemic regions, which largely superimpose *T. b. gambiense* infection areas<sup>1,9</sup>.

In western Africa, APOL1 C-terminal mutations that avoid SRA-mediated neutralization and allow humans to resist *T. b. rhodesiense* are very frequent<sup>28</sup>. This could have counter-selected *T. b. rhodesiense*, leaving room for *T. b. gambiense* parasites that resist APOL1 through a distinct mechanism. As this mechanism does not involve direct APOL1 neutralization it could allow parasite resistance to different APOL1 mutants or species-specific variants. However, *T. b. gambiense*

resistance to NHS can be bypassed under efficient APOL1 uptake (scheme in Supplementary Fig. 24), so that strategies targeting the parasite via the natural trypanolytic factor become conceivable.

## METHODS SUMMARY

The activity of TgsGP on trypanosome resistance to APOL1 was determined by evaluating relative trypanolysis by either human serum or recombinant APOL1 in assays combining purified recombinant (wild-type or mutant) TgsGP or synthetic peptides and various trypanosomes, either genetically transformed or not, as well as *in vitro* measurements on the effects of TgsGP on lipid membranes.

**Full Methods** and any associated references are available in the online version of the paper.

Received 13 December 2012; accepted 1 August 2013.

Published online 21 August 2013.

- Simarro, P. P. et al. The Atlas of human African trypanosomiasis: a contribution to global mapping of neglected tropical diseases. *Int. J. Health Geogr.* **9**, 57 (2010).
- Pays, E. et al. The trypanolytic factor of human serum. *Nature Rev. Microbiol.* **4**, 477–486 (2006).
- Raper, J., Fung, R., Ghiso, J., Nussenzweig, V. & Tomlinson, S. Characterization of a novel trypanosome lytic factor from human serum. *Infect. Immun.* **67**, 1910–1916 (1999).
- Vanhollebeke, B. & Pays, E. The trypanolytic factor of human serum: many ways to enter the parasite, a single way to kill. *Mol. Microbiol.* **76**, 806–814 (2010).
- Vanhollebeke, B. et al. A haptoglobin-hemoglobin receptor conveys innate immunity to *Trypanosoma brucei* in humans. *Science* **320**, 677–681 (2008).
- Vanhamme, L. et al. Apolipoprotein L-I is the trypanosome lytic factor of human serum. *Nature* **422**, 83–87 (2003).
- Pérez-Morga, D. et al. Apolipoprotein L-I promotes trypanosome lysis by forming pores in lysosomal membranes. *Science* **309**, 469–472 (2005).
- Berberof, M., Pérez-Morga, D. & Pays, E. A receptor-like flagellar pocket glycoprotein specific to *Trypanosoma brucei gambiense*. *Mol. Biochem. Parasitol.* **113**, 127–138 (2001).
- Rougemont, A. et al. Hypohaptoglobinaemia as an epidemiological and clinical indicator for malaria. Results of two studies in a hyperendemic region in West Africa. *Lancet* **332**, 709–712 (1988).
- Raper, J., Nussenzweig, V. & Tomlinson, S. The main lytic factor of *Trypanosoma brucei brucei* in normal human serum is not high density lipoprotein. *J. Exp. Med.* **183**, 1023–1029 (1996).

11. Gibson, W. Resolution of the species problem in African trypanosomes. *Int. J. Parasitol.* **37**, 829–838 (2007).
12. Xong, H. V. *et al.* A VSG expression site-associated gene confers resistance to human serum in *Trypanosoma rhodesiense*. *Cell* **95**, 839–846 (1998).
13. Lecordier, L. *et al.* C-terminal mutants of apolipoprotein L-I efficiently kill both *Trypanosoma brucei brucei* and *Trypanosoma brucei rhodesiense*. *PLoS Pathog.* **5**, e1000685 (2009).
14. Salmon, D. *et al.* A novel heterodimeric transferrin receptor encoded by a pair of VSG expression site-associated genes in *Trypanosoma brucei*. *Cell* **78**, 75–86 (1994).
15. Radwanska, M. *et al.* Novel primer sequences for a polymerase chain reaction-based detection of *Trypanosoma brucei gambiense*. *Am. J. Trop. Med. Hyg.* **67**, 289–295 (2002).
16. Felu, C., Pasture, J., Pays, E. & Pérez-Morga, D. Diagnosis potential of a conserved genomic rearrangement in the *Trypanosoma brucei gambiense*-specific TGS GP locus. *Am. J. Trop. Med. Hyg.* **76**, 922–929 (2007).
17. Gibson, W., Nemetschke, L. & Ndung'u, J. Conserved sequence of the TgsGP gene in Group 1 *Trypanosoma brucei gambiense*. *Infect. Genet. Evol.* **10**, 453–458 (2010).
18. Bussler, H., Linder, M., Linder, D. & Reinwald, E. Determination of the disulfide bonds within a B domain variant surface glycoprotein from *Trypanosoma congolense*. *J. Biol. Chem.* **273**, 32582–32586 (1998).
19. Blum, M. L. *et al.* A structural motif in the variant surface glycoproteins of *Trypanosoma brucei*. *Nature* **362**, 603–609 (1993).
20. Callebaut, I. *et al.* Deciphering protein sequence information through hydrophobic cluster analysis (HCA): current status and perspectives. *Cell. Mol. Life Sci.* **53**, 621–645 (1997).
21. Overath, P. & Engstler, M. Endocytosis, membrane recycling and sorting of GPI-anchored proteins: *Trypanosoma brucei* as a model system. *Mol. Microbiol.* **53**, 735–744 (2004).
22. Jean-François, F. *et al.* Aggregation of cateslytin  $\beta$ -sheets on negatively charged lipids promotes rigid membrane domains. A new mode of action for antimicrobial peptides? *Biochemistry* **47**, 6394–6402 (2008).
23. O'Brien, T. C. *et al.* A parasite cysteine protease is key to host protein degradation and iron acquisition. *J. Biol. Chem.* **283**, 28934–28943 (2008).
24. Kieft, R. *et al.* Mechanism of *Trypanosoma brucei gambiense* (group 1) resistance to human trypanosome lytic factor. *Proc. Natl Acad. Sci. USA* **107**, 16137–16141 (2010).
25. Higgins, M. K. *et al.* Structure of the trypanosome haptoglobin-hemoglobin receptor and implications for nutrient uptake and innate immunity. *Proc. Natl Acad. Sci. USA* **110**, 1905–1910 (2013).
26. DeJesus, E., Kieft, R., Albright, B., Stephens, N. A. & Hajduk, S. L. A single amino acid substitution in the group 1 *Trypanosoma brucei gambiense* haptoglobin-hemoglobin receptor abolishes TLF-1 binding. *PLoS Pathog.* **9**, e1003317 (2013).
27. Symula, R. E. *et al.* *Trypanosoma brucei gambiense* group 1 is distinguished by a unique amino acid substitution in the HpHb receptor implicated in human serum resistance. *PLoS Negl. Trop. Dis.* **6**, e1728 (2012).
28. Genovese, G. *et al.* Association of trypanolytic apoL1 variants with kidney disease in African-Americans. *Science* **329**, 841–845 (2010).
29. Douliez, J. P., Bellocq, A. M. & Dufourc, E. J. Effect of vesicle size, polydispersity and multilayering on solid-state P-31- and H-2-NMR spectra. *J. Chim. Phys.* **91**, 874–880 (1994).
30. Harrington, J. M. *et al.* Novel african trypanocidal agents: membrane rigidifying peptides. *PLoS ONE* **7**, e44384 (2012).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank M. Prévost for advice, C. Giroud and T. Baltz for help with *T. b. gambiense* culture adaptation, D. Horn for the gift of pTMF plasmid and E. Dupont for technical assistance. This work was supported by the Belgian Fund for Scientific Research, the Walloon WELBIO excellence programme, the Interuniversity Attraction Poles Programme–Belgian Science Policy, and the ERC grant 233312 TROJA. The CMMI is supported by the European Regional Development Fund and the Walloon Region. Financial support from the TGIR-RMN-THC Fr3050 (French high-field NMR network) and the Wellcome Trust is also acknowledged.

**Author Contributions** P.U. and E.P. conceived the work; P.U., S.U., L.L., F.F., P.T., F.H., A.G., V.Z., D.P.N., L.L., J.-M.C., A.P., C.F., P.P., B.V., S.K.M., J.L., J.S.P. and D.P.-M. performed experiments; J.C.M. provided *ICP* knockout parasites; P.U., E.J.D., D.P.-M. and E.P. supervised different aspects of the experimental plan; P.U. and E.P. wrote the paper.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to E.P. (epays@ulb.ac.be).



## METHODS

**Trypanosomes.** Most parasites analysed in this work have been described previously<sup>16</sup>. *ICP* knockout *T. b. brucei* parasites are described in ref. 31. The *T. b. gambiense* LiTat1.3 clone of the ELIANE strain was adapted to culture in IMDM supplemented medium as described<sup>32</sup>, except that we used 10% fetal calf serum (FCS, Sigma-Aldrich) and 10% Serum Plus (JRH Bioscience), and that the medium was incubated for 24 h at 37 °C in a CO<sub>2</sub>-equilibrated incubator before use. The parasites were either from culture or isolated by DE-52 column purification of infected rodent blood, as indicated. For high serum concentration experiments, the 100% normal human serum (NHS) or FCS medium was prepared after overnight dialysis of NHS or FCS against supplemented IMDM.

**Trypanolysis and *in vitro* growth assays.** Lysis assays were performed as described<sup>5</sup>. *In vitro* growth assays were performed by daily dilutions of trypanosomes at 10<sup>5</sup> per ml in IMDM-supplemented medium. Biological replicates were used. Normalizations were performed to reference situations. Identical results were obtained with *TgsGP* knockout parasites from the ELIANE (Ivory Coast) and BOSENDJA (Congo) strains<sup>16</sup>.

***In vivo* growth assays.** The animal experiments were IRB approved (local licence number LA1500474). 10<sup>4</sup> trypanosomes from tail blood in 100 µl PSGS medium were injected intraperitoneally in 8-week-old female NMRI mice. Daily, 2 µl tail blood samples were diluted in erythrocyte lysis buffer (0.85% (w/v) NH<sub>4</sub>Cl, 10 mM Tris pH 7.4) and counted with a haemocytometer. Sample size was empirically estimated and validated by the stability of the standard deviation of the growth curves. Within each experiment, animals used were morphologically indistinguishable (same strain, age and weight) and allocated randomly to each group. No blinding was done. Biological replicates were used.

***T. b. gambiense* transgenesis.** Stable transformations were obtained with the nucleofection method from Amaxa<sup>33</sup>. Briefly, 4 × 10<sup>7</sup> parasites were collected from culture of 5–8 × 10<sup>5</sup> cells per ml density, centrifuged and immediately resuspended in 100 µl of Amaxa Human T cell nucleofector solution (Lonza). Transfection of 20 µg of DNA was achieved in Nucleofector Device with program X-001. Transfected cells were re-suspended in 10 ml culture medium and incubated for 16 h before addition of suitable selection drugs and dilution to 100 ml in 24-well plates. A typical experiment gave rise to 10–30 positive wells out of 96 after 6–8 days.

***TgsGP* telomere deletion.** The pTMF plasmid vector<sup>34</sup> was modified for progressive deletion of the *TgsGP* sub-telomeric region. PCR amplicons of *Aut1* upstream region (primer set 5'-TTTCCCCGGGGTCTGCCTTTTCTCCAC TCACCTC-3' and 5'-TACGCGTCGACCTCTTCAAGTGTGGTGCCG-3'), *TgsGP* upstream region (primer set 5'-TTTCCCCGGGGCAGAGAAAGT GACGGAAGACG-3' and 5'-TACGCGTCGACTTAACAGAGAACTGCTACG CTTTGC-3') and *TgsGP* downstream region (primer set 5'-TTTCCCCGGGG GCGACAGCTGTTTATTATTCGCG-3' and 5'-TACGCGTCGACTAAATAA ATAACTGGCGCCGCTCG-3') were ligated into SmaI and SalI restriction sites of the pTMF vector. Plasmids were linearized with SmaI before transfection.

***TgsGP* knockout.** The unique *TgsGP* allele was replaced by the pPhleo-*TgsGP* knockout plasmid encoding resistance to phleomycin. pPhleo knockout was generated by PCR amplification of the phleomycin resistance gene and tubulin intergenic region from pBlueNTAP<sup>35</sup> using the primer set 5'-CCCCTCTAGACCAT GGCCAAGTTGACCAGTGCC-3' and 5'-CCCCAAGCTTGATCTGTGTCAGAA ATCAGACCGCG-3', and by ligation of this PCR product into the XbaI and HindIII restriction sites of pUC19. *TgsGP* upstream and downstream regions were PCR-amplified with the primer set 5'-TTTTCTCGAGTTGGTTCCATATC CCAATGTATAAGC-3' and 5'-TTTTCTAGATTCTCGTATTAGTTAGCTTT CTCTGC-3' and the primer set 5'-TTTTGGATCCCAAGCAATTCTTATG TGATGAAGG-3' and 5'-TTTTCTCGAGAAAATATTTCCAGACAATCAGA CAGTC-3', respectively. *TgsGP* 5' and 3' UTR amplicons were restricted with XhoI/XbaI and BamHI/XhoI respectively, and cloned into the BamHI/XbaI restriction sites of pPhleoKO. pPhleoKO was digested with XhoI before transfection.

***TgsGP* add-back and *TgsGP* mutants.** The pPARP promoter of pTSA-HYG2 (ref. 36) was removed by restriction using Ecl136II and PfuII, blunting with T4 polymerase supplemented with dTTP, and religation to generate pTSA-HYG2del. *TgsGP* cDNA was reverse-transcribed with primer (dT)<sub>18</sub>. PCR-amplified using the primer set 5'-CCCTCTAGAAGTACTCTCGAGTTTTTTTTTTTTTTT T-3' and 5'-GGTCGACATGTGGCAATTACTAGCAATAG-3', and restricted with SalI/XbaI before cloning into the SalI/SpeI restriction sites of pTSA-HYG2del, to generate the Tg19 plasmid. Site-directed mutagenesis of Tg19 was performed using the QuikChange Site-Directed Mutagenesis kit (Stratagene) following manufacturer instructions (primer sequences available on request). Tg19 was linearized with MluI before transfection of *T. b. gambiense TgsGP* knockout parasites, to generate add-back trypanosomes. Transfection of mutated versions of Tg19 into *T. b. gambiense TgsGP* knockout parasites generated Del- and Mut- parasites.

***In situ TgsGP* tagging.** The V5 epitope sequence was introduced into *TgsGP* immediately downstream from the signal peptide sequence. Tg19 was mutagenized to replace the ORF nucleotides 76–81 and 88–93, by the ApaI and AgeI restriction sites, respectively. The primer set 5'-CGGTAAGCCTATCCCT AACCTCTCTCGGTCTCGATTCTACGGGAGGCA-3' and 5'-CCGGTGC CTCCCGTAGAATCGAGACCGAGGAGAGGGTTAGGGATAGGCTTACCGG GCC-3' was annealed and cloned into ApaI- and AgeI- restricted mutant Tg19 to generate the Tg19-V5 plasmid. The *TgsGP* allele was modified by *in situ* insertion of pET-HYG-V5-*TgsGP* plasmid encoding resistance to hygromycin. The *TgsGP* 5' ORF including V5 epitope tag was PCR-amplified from Tg19-V5 plasmid with the primer set 5'-TTTTTAAGCTTATGTGGCAATTACTAGC AATAGC-3' and 5'-TTTCTCGAGCCGTCAATTGGAATTTGCT-3'. The *TgsGP* 5' UTR amplicon from the *TgsGP* knockout construct and *TgsGP* 5' V5 ORF amplicon were restricted with XhoI/XbaI and HindIII/XhoI, respectively, and cloned into the HindIII/XbaI restriction sites of pET-HYG-GFP<sup>37</sup>. The pET-HYG-V5-*TgsGP* plasmid was digested with XhoI before transfection.

***TbHpHbR* knockout.** Knockout of the *TbHpHbR* alleles of the *T. b. brucei* 'Single Marker' cell line<sup>38</sup> was performed with the pHygroKO-*HpHbR* and pPuroKO-*HpHbR* plasmids. The pHygroKO plasmid was generated by PCR amplification of the hygromycin resistance gene and intergenic region from the pHD328 vector<sup>38</sup> with the primer set 5'-TTTACATGTCCAAAAAGCCTGAACCTCACCGCG-3' and 5'-TTTTTAAGCTTAGGCCTTCGAGAATACTGCATAG-3' (KO-AS), restriction with HindIII/AflIII and cloning into HindIII/NcoI sites of pPhleoKO. The pPuroKO plasmid was generated as follow. The puromycin resistance gene was excised from pBS-Pur vector<sup>39</sup> with KpnI/EcoRV and cloned into MscI restricted pLew100 vector<sup>38</sup>. The puromycin resistance gene and 3' actin UTR were PCR-amplified with the primer set 5'-ATCACCATGGATCTG ACCGAG-3' and KO-AS, restricted with HindIII/NcoI and cloned into the corresponding sites of pPhleoKO to generate pPuroKO. Upstream and downstream regions of *TbHpHbR* were PCR-amplified respectively with the primer set 5'-TGCAGCCTCGAGCGTGTGAATGACATCAGCG-3' and 5'-GCTGCATC TAGACAAAGCTGCGACTGCACC-3', and the primer set 5'-TGCACAGG ATCCGACACCGTTTCTTCCAAAGACTGC-3' and 5'-TCTGCATCGAGC GGTGAGGCGCATTGTTCC-3'. *TbHpHbR* 5' and 3' UTR amplicons were restricted with XhoI/XbaI, and BamHI/XhoI respectively, and cloned into the BamHI/XbaI restriction sites of pHygroKO and pPuroKO. Plasmids were digested with XhoI before transfection.

***TbHpHbR* allele complementation.** PCR amplification of the *TbHpHbR* alleles from the *T. b. brucei* AnTat1.1E and *T. b. gambiense* LiTat1.3 genomic DNA, as well as cloning into the pTSARib vector, were performed as previously described<sup>5</sup>. BglII-restricted plasmids were transfected into *T. b. brucei HpHbR* knockout parasites<sup>5</sup> or *T. b. gambiense* LiTat1.3, as indicated.

***ICP* RNAi *T. b. brucei*.** A 260-bp fragment derived from the *ICP* open reading frame was amplified using the primer set 5'-TTTTGGATCCCGTATGG TCATTGGTGAAACC-3' and 5'-TTTTCTCGAGACGTGAATGTTGTAACG TTTGGC-3'. The amplification product was digested with BamHI and XhoI and ligated into the p2T7-177 plasmid<sup>38</sup> digested with the same enzyme mix. Linearized plasmid was transfected in single marker cell line<sup>38</sup>. RNAi was induced by addition of 1 µg ml<sup>-1</sup> doxycycline (Duchefa).

**Phylogenetic analyses.** Amino acid sequences of type B VSG N-terminal domains were retrieved from VSGdb ([http://leishman.cent.gla.ac.uk/kook\\_index.html](http://leishman.cent.gla.ac.uk/kook_index.html)) and aligned with *TgsGP* (CAJ87066.1) and SRA (AAC72381.1) sequences using Clustal Omega (<http://www.clustal.org/omega/>). The alignment was manually edited and curated with Gblock. Two independent runs were conducted in MrBayes until the standard deviation of split frequencies reached 0.05, sampling every 1,000 generations. Consensus tree excluded the first 25% of trees and was visualized with Archaeopteryx (<https://sites.google.com/site/cmzmasek/home/software/archaeopteryx>).

**qRT-PCR.** Quantitative RT-PCR was as described<sup>40</sup>. Briefly, isolated RNA was treated with DNase before the reverse transcription reaction, using TURBO DNase (Ambion) according to the manufacturer's instructions. The DNase was inactivated by addition of 0.1 volume of DNase inactivation buffer for 2 min at room temperature. Complementary DNA was synthesized with Transcriptor reverse transcriptase (Roche Applied Science) according to the manufacturer's instructions. The following primer sets were used for mRNA quantification: 5'-CGTCAGCAGCAAAGGTGTTA-3' and 5'-TGCATAAAGGACGAGTGCTG-3' for *TgsGP*, 5'-GAGGGGTCTTGTGTGGAAA-3' and 5'-ACTGGCATAA CTGCGGAAAC-3' for *HpHbR*, 5'-GGGGAATGGACGGTTACTTT-3' and 5'-GTGTTGGGTGCAAGAGGAAT-3' for cathepsin B, 5'-AGGGACCAAAAC CAATGTCTC-3' and 5'-AGTGGGACCTCCAACAACCTG-3' for cathepsin L, and 5'-CCCAGCAAAGTGACCTAATA-3' and 5'-CGGATGGTCTTGTGTGTT GTC-3' for inhibitor of cysteine peptidases (*ICP*) in Fig. 4b and 5'-CCGTAT GGTCAATGGTGAAACC-3' and 5'-ACGTGAATGTTGTAACGTTTGGC-3'.

for ICP in Supplementary Fig. 20. Normalization was performed with the primer set 5'-CACCGAAGTCTCCGTCAGT-3' and 5'-AGCTGAATTTCCCGTACA-3', targeting *H2B* mRNA. Technical replicates were applied.

**Recombinant proteins and peptides.** *TgsGP* and mutants were PCR-amplified from the Tg19 plasmid, or the corresponding mutant plasmid, with the primer set 5'-CCCACATGTCCATGGGTGCAGGCGAAAAATGGCGGCACGTAC-3' and 5'-CCCGTCGACCTCGAGTCCACCGCTTGCTGTGGTGTTCGCACTTCC C-3'. The amplicon was restricted with *NcoI* and *XhoI* and cloned into corresponding sites of pET21d (Novagen), fusing *TgsGP* ORF with six histidine residues. Purification was performed exactly as described<sup>5</sup>. Preparation of r-APOL1 was as described<sup>6</sup>, except that the pStaby1.2 system (Delphi Genetics) was used in this work. Synthetic peptides (>95% pure) were from Thermofisher or GenScript.

***TgsGP* knockout complementation with protein and peptides.** Parasites were incubated overnight with different amounts of peptides or NHS in 96-well plates. Each condition was assayed in triplicate. After room temperature equilibration, 20 µl of each lane was collected and mixed with 50 µl of CellTiter Glo Luminescent Cell Viability Assay buffer (Promega). Plates were mixed for 10 min before luminescent reading in Centro LB 960 (Berthold Tech) and resultant data were used to calculate parasite density by comparison with untreated controls. The linear relationship between the luminescence and the parasite density was checked for each tested populations as follows: the parasite concentration of serially diluted cultures was determined using a Neubauer slide and plotted against the luminescence values obtained from the Celltiter Glo assay. Technical replicates were applied.

**Protein labelling with Alexa Fluor.** Hp Alexa-488 labelling was described previously<sup>5</sup>. APOL1 labelling was performed after protein Ni-NTA elution (200 mM acetic acid, 0.4 M guanidium chloride) and dialysis, with Alexa Fluor 488 Antibody Labelling kit (Life Technologies) following the manufacturer's instructions. Alexa-633-labelled transferrin was purchased from Life Technologies.

**Antibodies.** Rat anti-*TgsGP* antibodies were obtained after subcutaneous immunization with r-*TgsGP* according to standard protocols. Antibodies were affinity purified on nitrocellulose membranes as follows: 1,000 µg of r-*TgsGP* was gel fractionated and transferred to Hybond-C Extra membranes (Amersham Biosciences). Ponceau-red-labelled band was excised and incubated overnight at 4 °C with 10 ml of 1:10 dilution of anti-*TgsGP* antibodies. After five extensive membrane washes in TBS, the antibodies were repeatedly eluted in 100 mM glycine (pH 2.7) and immediately neutralized with 0.1 M Tris (pH 8).

**Western blot analysis.** Western blots were incubated for 2 h with a 1:100 dilution of purified anti-*TgsGP* antibodies, 1:2,000 dilution of a rat polyclonal anti-APOL1 antibodies<sup>5</sup>, or with rat monoclonal anti-tubulin antibody (Abcam) in 150 mM NaCl, 0.5% (w/v) Tween 20, 20 mM Tris-HCl (pH 7.5) with 1% non-fat milk. The secondary antibodies, peroxidase-conjugated monoclonal mouse anti-rat IgGs (1:5,000; Serotec), were diluted in the same buffer and the bound antibodies were detected by chemiluminescence (Amersham). Albumin was visualized by Ponceau red staining of the membranes. The number of *TgsGP* molecules per cell was determined by three independent semiquantitative western blots (technical replicates), comparing cellular *TgsGP* signals to reference amounts of r-*TgsGP* by densitometry.

**Small-angle X-ray scattering analysis.** Small-angle X-ray scattering (SAXS) measurements were performed at the in-house lab-based instrument (Bruker AXS, prototype of Nanostar) at Aarhus University<sup>41</sup>. The instrument uses a rotating anode as source (Cu K $\alpha$ ) and a HiSTAR gas detector. The samples were in a thermo stated flow-through quartz capillary. A buffer solution was measured as background and water was used for absolute scale calibration. The data are given as a function of the scattering vector modulus  $q = 4\pi \sin(\theta)/\lambda$  where  $2\theta$  is the scattering angle and  $\lambda$  is the X-ray wavelength. Samples were measured at 4 °C at both high (2.4–5.6 mg ml<sup>-1</sup>) and low (0.6 mg ml<sup>-1</sup>) concentration. Technical replicates were applied.

**Immunofluorescence and flow cytometry.** PBS-washed cells were fixed in 2% paraformaldehyde for 10 min at 20 °C before being spread on poly-L-lysine-coated slides and subsequently treated with 0.1% (v/v) Triton X-100 in Tris-buffered saline for 10 min at 20 °C. The V5 epitope tag was detected by 1 h incubation with a 1:500 dilution of mouse monoclonal anti-V5 antibody (Life Technologies). Primary antibodies were detected with an Alexa-Fluor-594-conjugated goat anti-mouse IgG (Life Technologies). Cells were analysed with Zeiss LSM 710 confocal or Axioimager M2 epifluorescence microscope, as indicated. Flow cytometry analyses were performed exactly as described<sup>42</sup>. Technical replicates were applied.

**Electron microscopy measurement of endosomes.** Cells were fixed for 1 h at room temperature in 2.5% glutaraldehyde in culture medium, and postfixed in 2% OsO<sub>4</sub> in the same buffer. After serial dehydration in increasing ethanol concentrations, samples were embedded in agar 100 (Agar Scientific Ltd) and left to polymerize for 2 days at 60 °C. Ultrathin sections (50–70 nm thick) were collected in Formvar-carbon-coated copper grids by using a Leica EM UC6 ultramicrotome

and stained with uranyl acetate and lead citrate. Observations were made on a Tecnai 10 electron microscope (FEI), and images were captured with a MegaView II camera and processed with AnalySIS and Adobe Photoshop software. For each strain, 200 measurements were performed (technical replicates). D'Agostino-Pearson test assumed a non-Gaussian distribution and Fisher test validated the homogeneity of the variance.

**Immunogold labelling on trypanosome cryosections.** Cells (*TgsGP*-V5-tagged LiTat1.3) were fixed v/v in 2% paraformaldehyde, 0.2% glutaraldehyde in 0.1 M DPBS solution, embedded in 12% gelatin, 2.3 M sucrose and frozen in liquid nitrogen. Ultra-cryosectioning (80-nm cuts) was performed on a Leica EM UC7. Sections were collected on 100-mesh carbon formvar Ni-grids and were probed with a mouse monoclonal anti-V5 (Invitrogen), amplified by a biotinylated rabbit anti-mouse and detected by streptavidin 10-nm gold conjugate (BB International). Grids were mounted in methyl cellulose-1% uranyl acetate films. Observations were made on a Tecnai 10 electron microscope (FEI). The images were captured with a Olympus VELETA camera and processed with AnalySIS software (technical replicates).

#### Infrared spectroscopy of synthetic peptides in the presence of soy lipids.

Infrared spectra were recorded on an Equinox 55 spectrophotometer (Bruker Optics) equipped with a Golden Gate reflectance accessory (Specac). The internal reflection element was a diamond ATR crystal (2 × 2 mm) with an aperture angle of 45°. A total of 256 scans were accumulated for each spectrum. Spectra were recorded at a nominal resolution of 2 cm<sup>-1</sup>. Soy lipids (Avanti Polar Lipids) were dissolved in chloroform (0.5 mg ml<sup>-1</sup>). Dry lipid films were formed on a glass tube by slowly evaporating the solvent under a N<sub>2</sub> flux and dried overnight under vacuum. Peptides (0.5 mg) were dissolved 5 mg ml<sup>-1</sup> in 50% TFE/water (v/v). This solution was mixed with the dried lipid film to obtain a final peptide concentration of approximately 20% (w/w) and dried overnight under vacuum. This film was rehydrated in a 1 mM MES (pH 5.0), vortexed and sonicated. Samples (2 µl) were dried on the diamond crystal under a stream of N<sub>2</sub> and deuterated for 2 min before recording the spectra to remove the contribution of water. The amide I band (1,600–1,705 cm<sup>-1</sup>) which arises mainly from the stretching vibration of the carbonyl (C=O) from the peptide backbone was used for secondary structure determination<sup>43</sup>. Spectral intensities were normalized with respect to the amide I band area. To change the pH, the preparation was overlaid twice with 2 µl of 1 mM HEPES (pH 7.0) and incubated for 5 min. The buffer was removed with a pipette before the film was dried under a stream of N<sub>2</sub>. Technical replicates were applied.

**Isolation of lipid membranes from *T. b. brucei* bloodstream forms.** The method of ref. 44 was applied to MiTat1.1 *T. b. brucei* bloodstream forms grown in rats. The yield was 70.1 mg from 2.1 × 10<sup>11</sup> parasites.

**NMR analysis on DMPC/DMPG membranes.** A mixture of DMPC-<sup>2</sup>H<sub>54</sub>/DMPG (2:1 molar ratio) was dissolved in CHCl<sub>3</sub> and evaporated under nitrogen gas producing a lipid film. Lipids were purchased from Avanti Polar Lipids (Alabama). Peptides (lipid/peptide molar ratio of 30/1) were dissolved in TFE and added to the lipid film. The solvent was evaporated and the residue hydrated in 50 ml of water, shaken and lyophilized overnight. The fluffy powder was rehydrated (80%, w/w) in either 1 mM MES (pH 5) or 10 mM Tris, 10 mM KCl, 0.5 mM EDTA (pH 7) to reach a lipid concentration of 0.3 M. All preparations were subjected to three freeze-thaw cycles and shaking in vortex mixer for better homogenization.

NMR experiments were carried out at 107.4 MHz for deuterium and 283.4 MHz for phosphorus, on a Bruker Avance III 700 SB spectrometer with a CP-MAS triple 4 mm <sup>1</sup>H/<sup>2</sup>H/<sup>31</sup>P probe. Echo sequences were performed to record time-dependent signals that were Fourier transformed with 100–200 Hz Lorentzian filtering to yield wide line spectra<sup>45</sup>. Pulse durations and separations were respectively 5 and 30–35 µs, spectral windows were 250–1,000 kHz and 1K acquisitions were accumulated with repetition rates of 1.5 to 5 s for proper return to magnetic equilibrium after each scan. Experiments were performed in the 10–45 °C range allowing 20 min equilibrium time at each temperature before NMR sampling. Spectral calculations and simulations were performed to obtain lipid ordering and orientation. Technical replicates were applied.

**NMR analysis on trypanosome lipid membranes.** 9 mg of lipids isolated from trypanosome membranes were co-solubilized in CHCl<sub>3</sub> with 1 mg of <sup>2</sup>H<sub>31</sub>-palmitoylcholinephosphatidylcholine (Avanti Polar Lipids) and evaporated under nitrogen gas producing a lipid film. Addition of peptides was performed as described for synthetic lipids. NMR experiments were carried out at 122.8 MHz for deuterium, on a Bruker Avance III 800 SB spectrometer with a CP-MAS dual 4 mm <sup>1</sup>H/<sup>2</sup>H DVT probe. Acquisition was performed as described for synthetic lipids except that the number of scans was increased to 18k. Technical replicates were applied.

**FRAP analysis.** FRAP assays were performed essentially as described<sup>30</sup>. Briefly, 2 × 10<sup>6</sup> bloodstream *T. brucei* cells (single marker line<sup>38</sup>) were labelled with 1 µM sulpho-NHS coupled Atto 488 fluorescent dye (ATTO-TEC GmbH, Siegen) for 15 min on ice and washed 2 times with cold phosphate-saline buffer. Cells were then

incubated for 15 min at 37 °C with different concentrations of TgsGP inter-helical wild-type or mutant peptides (stock 10 mM in DMSO), using incubations with the same volumes of DMSO as controls. The trypanosomes were immobilized by embedding in 12% gelatin-buffer solution and mounted on a 1% low melting point agarose pad sealed with rubber glue. The samples were put in an incubation chamber thermo-stabilized at 30 °C. Under these conditions the parasites remained alive for at least 6 h. FRAP analysis was performed with a Zeiss LSM 710 confocal microscope. For each cell, 20 pre-bleach and 100 post-bleach acquisitions were made with 500-ms intervals. 40 cells were analysed per condition. The mobile fraction was calculated according to the Zeiss Zen 2010 software using a double normalization (background and reference image). Replicate number was determined empirically to 40 by collecting data until the standard error of the mean value was stable (technical replicates). The Gaussian distribution of the data and the homogeneity of the variance were evaluated by the D'Agostino–Pearson test and Fischer test, respectively.

**Surface plasmon resonance analysis.** Surface plasmon resonance (SPR) analysis was conducted on a Biacore 3000 instrument (GE Healthcare), essentially as in ref. 5. The receptor variants were immobilized in 10 mM sodium acetate (pH 4.5), and remaining sites were blocked with 1 M ethanolamine (pH 8.5). The SPR signal generated from the immobilized receptors corresponded to 0.09–0.1 pmol of protein per mm<sup>2</sup>. For binding analyses the ligands were dissolved in running buffer (10 mM HEPES, 150 mM NaCl, 3.0 mM CaCl<sub>2</sub>, 1.0 mM EGTA, 0.005% Nonidet P20, pH 7.4). In contrast to previous SPR experiments<sup>5</sup>, recordings with r-HPR showed a weak binding signal of approximately 100 response units in the absence of Hb. This most probably represents unspecific binding of the recombinant material as a similar response was observed with uncomplexed r-Hp, but not with uncomplexed native Hp from NHS. Purified human Hb A0 and Hp (phenotype 1-1) were from Sigma. Technical replicates were applied.

**Measurement of endosomal pH.** The accumulation of the weak base [<sup>14</sup>C]methylamine was used to investigate the intracellular pH and the pH of acidic organelles in *T. b. gambiense* as described previously for *T. b. brucei*<sup>46</sup>. Cells isolated in the ascending phase of rat parasitaemia ( $2-3 \times 10^7$  ml<sup>-1</sup>) were incubated at 37 °C in TES buffer (pH 7.5) in the presence of [<sup>14</sup>C]methylamine (0.1 µCi ml<sup>-1</sup>; 1.8 µM) for 40 min. The cells were separated from the medium by rapid centrifugation through a 200 µl oil layer (2:1, v/v, mixture of di-*N*-butylphthalate and di-*iso*-octylphthalate). A parallel experiment was performed to determine the intracellular volume and the amount of trapped extracellular probe present in the cell pellet by incubation with <sup>3</sup>H<sub>2</sub>O (1 µCi ml<sup>-1</sup>) and [<sup>14</sup>C]carboxyininulin (0.1 µCi ml<sup>-1</sup>). The accumulation ratio was determined by dividing the intracellular concentration of the probe by the extracellular concentration. The cytoplasmic pH (pH<sub>c</sub>) was assessed by measuring methylamine accumulation in the presence of chloroquine (0.3 mM) or bafilomycin (1.8 µM). Comparison of the endosome size between *T. b. brucei* and *T. b. gambiense* was performed by measurement of the surface of endosomal vesicle sections by transmission electron microscopy. Biological replicates were used.

**Detection of cysteine protease activity.** The zymogram analysis was performed as in ref. 47. Briefly, cells were re-suspended at 10<sup>8</sup> cells per ml in sample buffer (25 mM BisTris, 25 mM acetate, 2% (w/v) LDS, 15% (v/v) glycerol, 0.05% (w/v) bromophenol blue, 0.05% (w/v) phenol red) and stored at -20 °C. Separation gels (10.7% (w/v) acrylamide, 0.3% (w/v) bisacrylamide, 10% (v/v) glycerol, 0.1% (w/v) gelatin, 0.3 M BisTris, and 0.3 M acetate, pH 5.67, 1% (w/v) riboflavin) and

stacking gels (3.9% (w/v) acrylamide, 0.1% (w/v) bisacrylamide, 2% (v/v) glycerol, 0.3 M BisTris, and 0.3 M acetate, pH 5.67, 1% (w/v) riboflavin) were polymerized by 1 h exposure to neon light. Migration was performed in cold room at pH 5.8 with 0.14 M MES, 56 mM BisTris, 0.1% (w/v) LDS (cathode buffer) and 0.1 M BisTris, 0.1 M MES (anode buffer). The gels were washed three times for 10 min in 2.5% (v/v) Triton X-100, 50 mM sodium acetate, 100 mM sodium chloride, 10 mM cysteine, pH 5.5 and incubated for 16 h at 37 °C in 50 mM sodium acetate, 100 mM sodium chloride, 10 mM cysteine, 5 mM EDTA, pH 5.5 before Coomassie staining. Biological replicates were used.

31. Santos, C. C., Coombs, G. H., Lima, A. P. & Mottram, J. C. Role of the *Trypanosoma brucei* natural cysteine peptidase inhibitor ICP in differentiation and virulence. *Mol. Microbiol.* **66**, 991–1002 (2007).
32. Baltz, T., Baltz, D., Giroud, C. & Crockett, J. Cultivation in a semi-defined medium of animal infective forms of *Trypanosoma brucei*, *T. equiperdum*, *T. evansi*, *T. rhodesiense* and *T. gambiense*. *EMBO J.* **4**, 1273–1277 (1985).
33. Burkard, G., Fragoso, C. M. & Roditi, I. Highly efficient stable transformation of bloodstream forms of *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* **153**, 220–223 (2007).
34. Glover, L., Alsford, S., Beattie, C. & Horn, D. Deletion of a trypanosome telomere leads to loss of silencing and progressive loss of terminal DNA in the absence of cell cycle arrest. *Nucl. Acids Res.* **35**, 872–880 (2007).
35. Lecordier, L. et al. Characterization of a TFIIF homologue from *Trypanosoma brucei*. *Mol. Microbiol.* **64**, 1164–1181 (2007).
36. Sommers, J., Peterson, G., Keller, G. A., Parsons, M. & Wang, C. C. The C-terminal tripeptide of glycosomal phosphoglycerate kinase is both necessary and sufficient for import into the glycosomes of *Trypanosoma brucei*. *FEBS Lett.* **316**, 53–58 (1993).
37. Devaux, S. et al. Diversification of function by different isoforms of conventionally shared RNA polymerase subunits. *Mol. Biol. Cell* **18**, 1293–1301 (2007).
38. Wirtz, E., Leal, S., Ochatt, C. & Cross, G. A. A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* **99**, 89–101 (1999).
39. Ruepp, S. et al. Survival of *Trypanosoma brucei* in the tsetse fly is enhanced by the expression of specific forms of procyclin. *J. Cell Biol.* **137**, 1369–1379 (1997).
40. Salmon, D. et al. Cytokinesis of *Trypanosoma brucei* bloodstream forms depends on expression of adenyl cyclases of the ESAG4 or ESAG4-like subfamily. *Mol. Microbiol.* **84**, 225–242 (2012).
41. Pedersen, J. S. A flux- and background-optimized version of the NanoSTAR small-angle X-ray scattering camera for solution scattering. *J. Appl. Cryst.* **37**, 369–380 (2004).
42. Vanhollebeke, B., Uzureau, P., Monteyne, D., Pérez-Morga, D. & Pays, E. Cellular and molecular remodelling of the endocytic pathway during differentiation of *Trypanosoma brucei* bloodstream forms. *Euk. Cell* **9**, 1272–1282 (2010).
43. Goormaghtigh, E., Cabiaux, V. & Ruyschaert, J. M. Determination of soluble and membrane protein structure by Fourier transform infrared spectroscopy. III. Secondary structures. *Subcell. Biochem.* **23**, 405–450 (1994).
44. Folch, J., Lees, M. & Sloane Stanley, G. H. A simple method for the isolation and purification of total lipids from animal tissues. *J. Biol. Chem.* **226**, 497–509 (1957).
45. Sani, M. A., Castano, S., Dufourc, E. J. & Gröbner, G. Restriction of lipid motion in membranes triggered by β-sheet aggregation of the anti-apoptotic BH4 domain. *FEBS J.* **275**, 561–572 (2008).
46. Nolan, D. P. & Voorheis, H. P. Hydrogen ion gradients across the mitochondrial, endosomal and plasma membranes in bloodstream forms of *Trypanosoma brucei*. *Eur. J. Biochem.* **267**, 4601–4614 (2000).
47. Klose, A., Zigrino, P., Denhöfer, R., Mauch, C. & Hunzelmann, N. Identification and discrimination of extracellularly active cathepsins B and L in high-invasive melanoma cells. *Anal. Biochem.* **353**, 57–62 (2006).



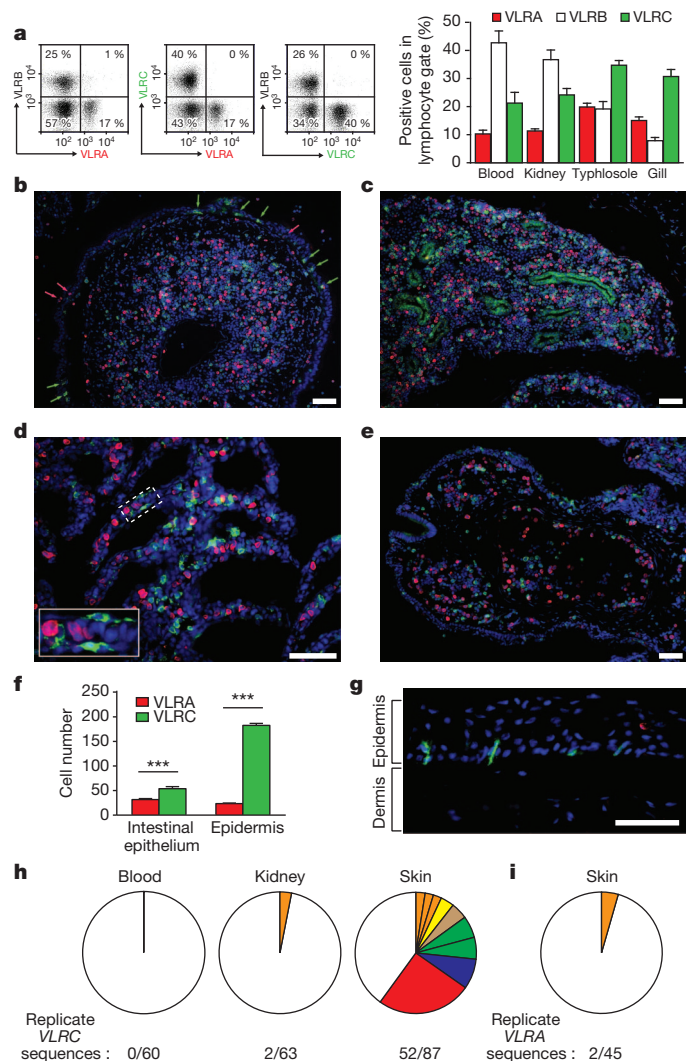
# Evolutionary implications of a third lymphocyte lineage in lampreys

Masayuki Hirano<sup>1</sup>, Peng Guo<sup>1†</sup>, Nathanael McCurley<sup>1</sup>, Michael Schorpp<sup>2</sup>, Sabyasachi Das<sup>1</sup>, Thomas Boehm<sup>2</sup> & Max D. Cooper<sup>1</sup>

Jawed vertebrates (gnathostomes) and jawless vertebrates (cyclostomes) have different adaptive immune systems<sup>1,2</sup>. Gnathostomes use T- and B-cell antigen receptors belonging to the immunoglobulin superfamily<sup>3,4</sup>. Cyclostomes, the lampreys and hagfish, instead use leucine-rich repeat proteins to construct variable lymphocyte receptors (VLRs), two types of which, VLRA and VLRB, are reciprocally expressed by lymphocytes resembling gnathostome T and B cells<sup>5–7</sup>. Here we define another lineage of T-cell-like lymphocytes that express the recently identified VLRC receptors<sup>8,9</sup>. Both VLRC<sup>+</sup> and VLRA<sup>+</sup> lymphocytes express orthologues of genes that gnathostome  $\gamma\delta$  and  $\alpha\beta$  T cells use for their differentiation, undergo VLRC and VLRA assembly and repertoire diversification in the ‘thymoid’ gill region, and express their VLRs solely as cell-surface proteins. Our findings suggest that the genetic programmes for two primordial T-cell lineages and a prototypic B-cell lineage were already present in the last common vertebrate ancestor approximately 500 million years ago. We propose that functional specialization of distinct T-cell-like lineages was an ancient feature of a primordial immune system.

The invariant stalk region of the sea lamprey VLRC shares approximately 20% sequence identity with the invariant VLRA and VLRB stalk regions, and distinguishing VLRC sequences are also present in the amino-terminal and carboxy-terminal leucine-rich repeat regions<sup>8,9</sup>. Four mouse monoclonal antibodies specific for the VLRC protein were produced, all of which identified a third lymphocyte population in lampreys that did not express VLRA or VLRB (Fig. 1a and Supplementary Fig. 1). VLRC<sup>+</sup> lymphocytes were more numerous than VLRA<sup>+</sup> lymphocytes in the principal lymphoid tissues of lamprey larvae (blood, kidneys, typhlosole and gill region) and they constituted the majority of lymphocytes in typhlosole and gills, whereas VLRB<sup>+</sup> lymphocytes predominated in blood and kidneys (Fig. 1a).

Immunofluorescence analysis of tissue sections indicated a similar distribution pattern for VLRA<sup>+</sup> and VLRC<sup>+</sup> lymphocytes in the typhlosole (Fig. 1b), kidneys (Fig. 1c), gills (Fig. 1d) and hypopharyngeal fold (Fig. 1e). Both cell types were round or oval in shape within blood vessels and interstitial spaces of the kidneys and typhlosole, but were dendritic in shape in the gill and intestinal epithelium (see inset in Fig. 1d) where the VLRC<sup>+</sup> cells were more numerous (VLRC/VLRA ratio of approximately 1.7/1) (Fig. 1b, f) and VLRB<sup>+</sup> cells were infrequent. VLRC<sup>+</sup> cells with inter-digitating morphology were the dominant lymphocyte type in the epidermal region of the skin (VLRC/VLRA ratio of 8/1) (Fig. 1f, g), a site in which VLRB<sup>+</sup> cells were rarely observed. The predominance of VLRC<sup>+</sup> cells in the epidermis is reminiscent of the dendritic epidermal T cells in mice, which express the same canonical  $\gamma\delta$  T-cell receptor (TCR)<sup>10</sup>. When VLRC sequences for the VLRC<sup>+</sup> cells in different tissues were compared, repetitive VLRC sequences were abundant in the skin, but rare in kidney and blood samples (Fig. 1h). We found several examples of identical or almost identical VLRC sequences in skin samples from different animals (Supplementary Fig. 2), suggesting that the VLRC repertoire in the skin is less diverse and more stereotypic than elsewhere. Importantly,



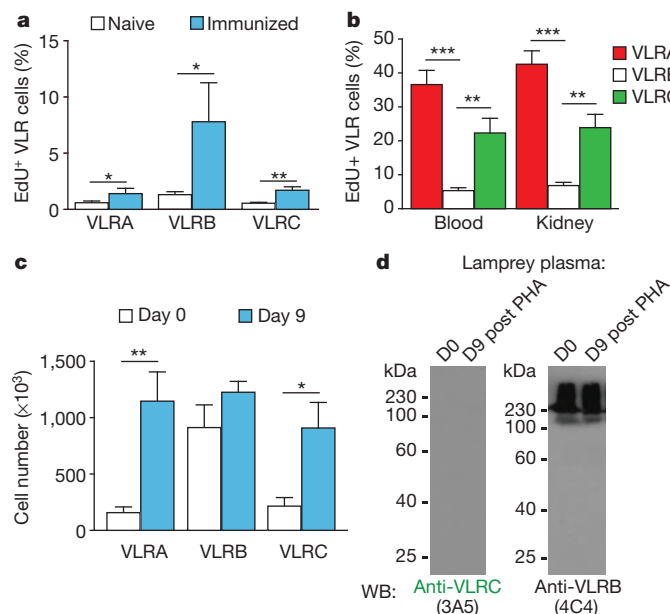
**Figure 1 | Tissue distribution of VLRA<sup>+</sup>, VLRB<sup>+</sup> and VLRC<sup>+</sup> lymphocytes.** **a**, Flow cytometric analysis of lymphocyte-gated cells stained with monoclonal antibodies specific for VLRA (R110), VLRB (4C4) and VLRC (3A5) (left). Lymphocyte population in lamprey larvae (right);  $n = 11$ . **b–g**, Immunofluorescence staining of VLRC<sup>+</sup> (green) and VLRA<sup>+</sup> (red) lymphocytes in larval tissue sections, and DAPI counterstaining of nuclei (blue); scale bars, 50  $\mu$ m (**b–e**, **g**). Shown are typhlosole surrounded by intestinal epithelium (intraepithelial lymphocytes) (**b**, arrows), kidneys (**c**), gill filaments (**d**; inset is a magnification of the area indicated by the dashed box), hypopharyngeal fold (**e**), skin (**g**). Lymphocyte distribution in the intestinal epithelium and skin epidermis is shown in **f**;  $n = 5$  larvae. \*\*\* $P < 0.0001$ ; error bars, s.e.m. **h**, Frequency of replicate VLRC sequences in indicated tissues from two larvae; replicate VLRC sequences were not shared by different tissues. **i**, Frequency of VLRA replicates. The numbers of clonal replicates are colour-coded: red, 22; blue, 7; green, 5; pale orange, 4; yellow, 3; orange, 2.

<sup>1</sup>Emory Vaccine Center and Department of Pathology and Laboratory Medicine, Emory University, 1462 Clifton Road North-East, Atlanta, Georgia 30322, USA. <sup>2</sup>Department of Developmental Immunology, Max Planck Institute of Immunobiology and Epigenetics, Stuebeweg 51, Freiburg D-79108, Germany. <sup>†</sup>Present address: Novo Nordisk Research Centre China, 20 Life Science Park Road, Beijing 102206, China.

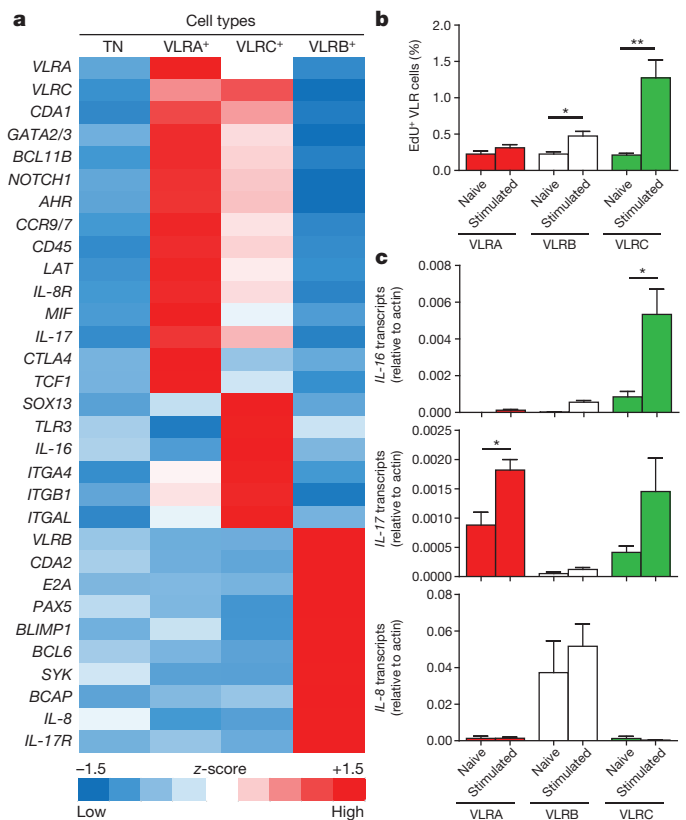
restricted diversity was not observed for VLRA sequences isolated from the same skin samples (Fig. 1i).

Antigen-binding VLRB<sup>+</sup> lymphocytes respond to immunization with proliferation and differentiation into plasma cells that secrete VLRC antibodies<sup>11,12</sup>. VLRA<sup>+</sup> lymphocytes also proliferate in response to immunization, but fail to bind unmodified immunogens, either before or after immunization, and do not differentiate into VLRA-secreting cells<sup>5</sup>. Proliferative responses to *Bacillus anthracis* exosporium were also observed for VLRC<sup>+</sup> lymphocytes (Fig. 2a and Supplementary Fig. 3a). The VLRC<sup>+</sup> lymphocytes also responded to the plant mitogen phytohaemagglutinin (PHA) with vigorous proliferation (Fig. 2b) and increased cell numbers (Fig. 2c and Supplementary Fig. 3b) and the activated VLRC<sup>+</sup> cells resembled activated VLRA<sup>+</sup> cells, in that they were large lymphoblasts with limited endoplasmic reticulum (Supplementary Fig. 3c). Nevertheless, plasma samples from naive or PHA-stimulated lampreys were devoid of VLRC protein (Fig. 2d) and VLRC transfectants did not secrete the VLRC protein (Supplementary Fig. 4).

Gene-expression profiles were compared for VLRA<sup>+</sup>, VLRB<sup>+</sup>, VLRC<sup>+</sup> and triple-negative populations of cells with lymphocyte light-scattering characteristics by examining the expression levels of a selected panel of orthologous genes. Discriminating profiles that were observed for the different populations included genes for transcription factors, cytokines or chemokines and their receptors, integrins, Toll-like receptors (TLRs), and various signalling molecules (Fig. 3a and Supplementary Tables 1 and 2). This analysis confirmed the clear dichotomy of the VLRB<sup>+</sup> B-like and the VLRA<sup>+</sup> T-cell-like lymphocytes. Although the transcriptional profiles were more similar for VLRA<sup>+</sup> and VLRC<sup>+</sup> cells, they differed in certain aspects; genes were expressed preferentially by VLRA<sup>+</sup> lymphocytes for the T-cell factor 1 (TCF1, also known as transcription factor 7), a key transcription factor for  $\alpha\beta$  T-cell lineage determination, and CTLA4, an important regulatory co-receptor for mammalian T cells. Conversely, VLRC<sup>+</sup> lymphocytes preferentially expressed the SRY-box containing gene 13 (SOX13) encoding a fate-determining transcription factor used for  $\gamma\delta$  T-cell lineage commitment<sup>13,14</sup>, an integrin  $\alpha$ L (ITGAL) orthologue of one component of the heterodimeric lymphocyte function associated antigen 1 (LFA1), and integrins



**Figure 2 | Antigen and mitogen responses.** **a**, Lymphocyte proliferation in typhlosol before ( $n = 6$ ) and 28 days after ( $n = 3$ ) *B. anthracis* exosporium immunization measured by EdU (5-ethynyl-2'-deoxyuridine) incorporation. **b**, Lymphocyte proliferation 9 days after PHA stimulation ( $n = 7$ ). **c**, Lymphocyte numbers in blood after PHA stimulation ( $n = 7$ ). **d**, Western blot (WB) analysis of plasma before (day 0 (D0)) and 9 days after (D9) PHA stimulation. kDa, kilodaltons. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ ; error bars, s.e.m.

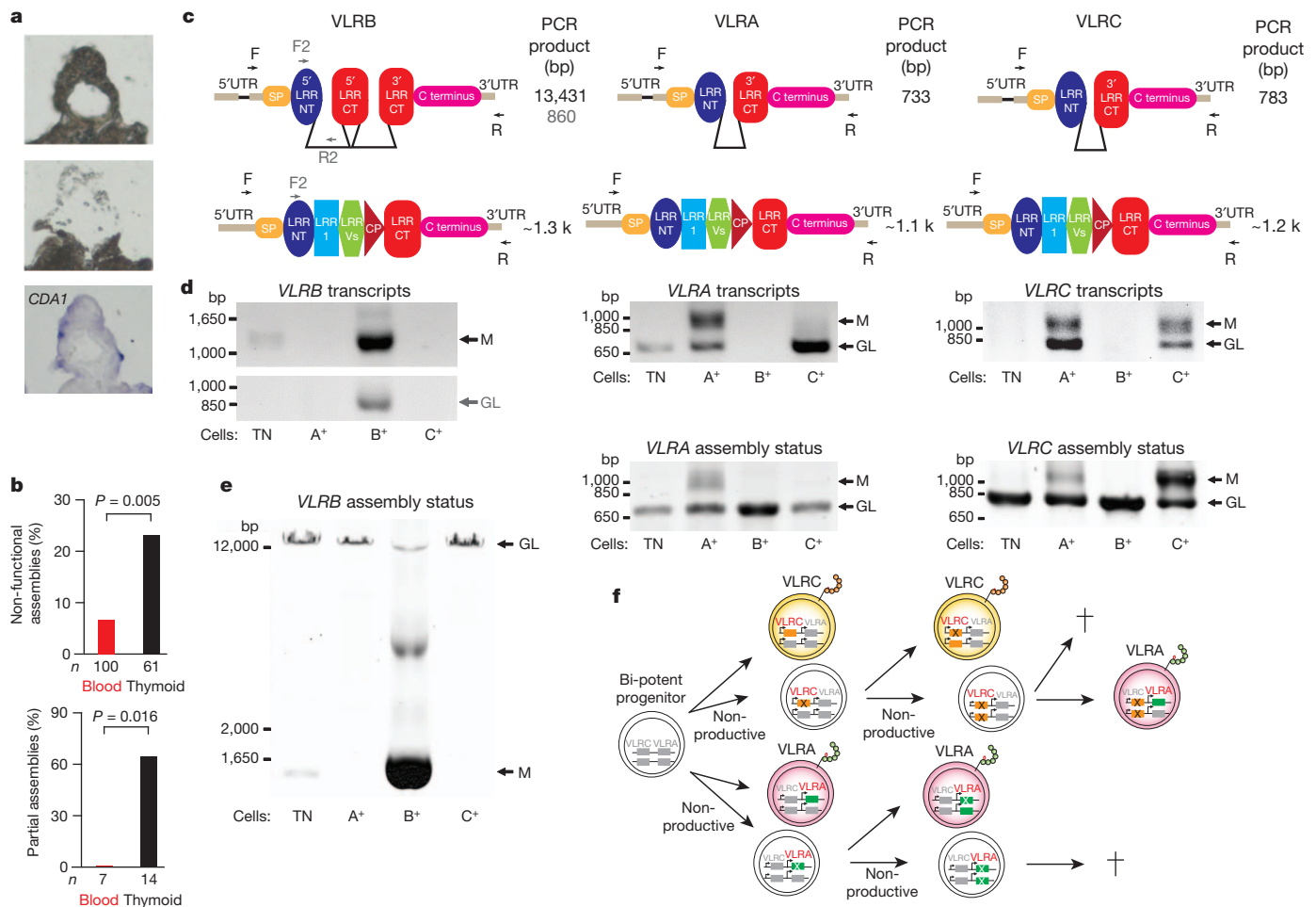


**Figure 3 | Gene-expression profiles of VLRA<sup>+</sup>, VLRB<sup>+</sup> and VLRC<sup>+</sup> lymphocytes and their poly(I:C) responses.** **a**, Relative transcript levels of the indicated genes were measured by quantitative PCR for purified VLRA<sup>+</sup>, VLRB<sup>+</sup>, VLRC<sup>+</sup> and triple negative (TN) lymphocyte populations and compiled into a heat map ( $n = 5$ ). **b**, Proliferative responses to *in vivo* poly(I:C) stimulation, measured by EdU incorporation ( $n = 3$ ). **c–e**, Cytokine expression of purified blood lymphocytes before and 9 days after poly(I:C) treatment of lamprey larvae measured by quantitative PCR ( $n = 3$ ): **c**, IL-16, **d**, IL-17 and **e**, IL-8. \* $P < 0.05$ , \*\* $P < 0.01$ ; error bars, s.e.m.

$\alpha 4$  and  $\beta 1$  (ITGA4 and ITGB1) orthologues of the two components of very late antigen 4 (VLA4), the expression of which was correlated with adherence of human  $\gamma\delta$  T cells to epithelial cells<sup>15</sup>, TLR3 (ref. 16) and interleukin-16 (IL-16), a modulator of T-cell activation<sup>17</sup>.

The functional implication of the preferential TLR3 expression by lamprey VLRC<sup>+</sup> lymphocytes was examined by comparing the VLRA<sup>+</sup>, VLRB<sup>+</sup> and VLRC<sup>+</sup> lymphocyte responses to poly(I:C), a structural mimic of double-stranded RNA. This analysis revealed a preferential proliferative response by VLRC<sup>+</sup> lymphocytes (Fig. 3b); VLRB<sup>+</sup> lymphocytes also responded to a lesser extent. Poly(I:C) stimulation also enhanced the expression of IL-16 transcripts by VLRC<sup>+</sup> cells, while having only a slight effect on IL-17 expression by the VLRA<sup>+</sup> and VLRB<sup>+</sup> cells and no effect on IL-8 expression by VLRB<sup>+</sup> cells (Fig. 3c). These results suggest that TLR3 expression by VLRC<sup>+</sup> lymphocytes could potentially facilitate their participation in responding to infection with RNA viruses, akin to the modulation of  $\gamma\delta$  T cell responses by TLRs in mammals<sup>18</sup>.

To determine whether VLRC assembly accompanies VLRA assembly in the thymus-equivalent region at the tips of the gill filaments<sup>19</sup>, we compared DNA samples obtained by laser capture micro-dissection of cells from the 'thymoid' region, which is characterized by ongoing CDA1 expression, and from blood where CDA1 expression is undetectable (Fig. 4a, b). Functional and non-functional VLRC assemblies were observed in both locations, although non-functional VLRC assemblies were present in higher frequency in the 'thymoid' region (14 non-functional out of 61 total distinct VLRC assemblies) than in blood (7 non-functional out of 100 total VLRC assemblies). Moreover, most of the non-functional



**Figure 4 | Analysis of *VLRC*, *VLRA* and *VLRB* transcription and assembly.** **a**, Thymoid procurement site before (top) and after (middle) laser-capture micro-dissection; *CDA1*-expressing cells (blue) were detected by RNA *in situ* hybridization in an adjacent section (bottom). **b**, Proportion of non-productive sequences among assembled *VLRC* genes (top) and partially assembled genes among non-productive *VLRC* sequences (bottom). **c**, Schematic of *VLRB*, *VLRA* and *VLRC* genes before (top) and after (bottom) assembly. Forward (F) and reverse (R) primer locations and predicted PCR product sizes are indicated. The F2 and R2 primer pair (grey) was used to amplify *VLRB* transcripts. **d**, Total RNA extracted from purified lymphocyte populations was amplified by RT-PCR

(PCR with reverse transcription). bp, base pairs; GL, germline transcripts; M, transcripts of assembled *VLR* genes. **e**, Assembly of *VLRB*, *VLRA* and *VLRC* genes identified by PCR of genomic DNA. **f**, Schematic model of *VLRA* and *VLRC* assembly during the development of bi-potent precursor cells in the 'thymoid'. Orange bars indicate *VLRC* assembly, green bars indicate *VLRA* assembly (solid coloured bars, productive assembly; bars with a cross, non-productive assembly). Green and orange chains represent leucine-rich repeat modules of *VLRA* and *VLRC*, respectively. Red text represents genes that are assembled. Grey text represents genes that are not assembled. Dagger symbols represent cell death.

*VLRC* sequences in cells within the 'thymoid' region consisted of incomplete intermediate-stage assemblies (9 out of 14), whereas incomplete *VLRC* assemblies were not found among the non-functional *VLRC* assemblies in blood (0 out of 7).

Transcription profiles for the *VLR* loci and their assembly status were examined in *VLRA*<sup>+</sup>, *VLRB*<sup>+</sup>, *VLRC*<sup>+</sup> and triple-negative lymphocyte populations purified by fluorescence-activated cell sorting from pooled cell suspensions derived from the blood, gills, kidneys, typhlosole and skin of lamprey larvae. Each sample was divided into two aliquots, which were used for genomic DNA and RNA transcript analysis. The polymerase chain reaction (PCR) products for the assembled or mature *VLRA* and *VLRC* genes were larger than those for the germline genes (see Fig. 4c), thus allowing their discrimination, purification and sequence analysis. Both germline and assembled *VLRB* genes were transcribed in *VLRB*<sup>+</sup> cells, whose *VLRA* and *VLRC* genes were transcriptionally silent (Fig. 4d). Conversely, *VLRB* transcription was not detected in *VLRA*<sup>+</sup> or *VLRC*<sup>+</sup> cells. However, both *VLRA* and *VLRC* loci were transcriptionally active in *VLRA*<sup>+</sup> and *VLRC*<sup>+</sup> lymphocytes. Hence, the expression of *VLRA* and *VLRC* genes was not lineage-specific, in contrast to the distinct cell surface phenotypes of *VLRA*<sup>+</sup> and *VLRC*<sup>+</sup> lymphocytes.

Consistent with previous studies indicating that mono-allelic assembly is the rule for *VLRA*<sup>+</sup> lymphocytes<sup>5</sup>, we observed an approximately 50/50 ratio of assembled versus non-assembled *VLRA* genes in *VLRA*<sup>+</sup> cells (Fig. 4e) and sequence analysis indicated that 94% (99 out of 105) of the *VLRA* assemblies in *VLRA*<sup>+</sup> cells were functional. *VLRC* assemblies were also detected in the *VLRA*<sup>+</sup> cells (Fig. 4e), but 79% (46 out of 58) of these were non-productive assemblies. The *VLRA* and *VLRC* assembly status was more complex for the *VLRC*<sup>+</sup> cells, as indicated by the 65/35 ratio of assembled versus non-assembled *VLRC* genes (Fig. 4d, e). The predominance of assembled gene products in *VLRC*<sup>+</sup> cells suggested the presence of bi-allelic *VLRC* assemblies in a fraction of these cells, and approximately 15% (20 out of 132) of the *VLRC* assemblies in *VLRC*<sup>+</sup> cells were non-productive. Conversely, *VLRA* assemblies were rarely detected in *VLRC*<sup>+</sup> cells, and almost all of these were non-productive (54 out of 55). Most non-functional *VLRA* and *VLRC* sequences contained frame-shift mutations (Supplementary Fig. 5), and stop-codon mutations were less common. These composite results suggest a differentiation model in which bi-potential precursor cells begin to undergo *VLRC* or *VLRA* assembly in the 'thymoid', and the assembly process may proceed to involve both alleles and even both loci to achieve a



productive VLRC or VLRA assembly in order to survive as receptor positive cells (Fig. 4f).

The present studies demonstrate that lampreys possess two T-cell-like lineages that exhibit mutually exclusive expression of their VLRA and VLRC surface proteins and have overlapping but distinct gene expression profiles. Their currently defined characteristics are reminiscent of the two principal T-cell lineages in jawed vertebrates. In several aspects, the gene-expression profiles of VLRA<sup>+</sup> and VLRC<sup>+</sup> cells resemble those of mammalian TCRαβ<sup>+</sup> and TCRγδ<sup>+</sup> T cells, although these lymphocyte lineages clearly cannot be considered to be precise equivalents. In addition to the convergent evolution of entirely different types of recognition receptors, other divergent modifications have almost certainly occurred during the approximately 500 million years of evolution since jawless and jawed vertebrates last shared a common ancestor. It will thus be particularly important to establish the degree of functional similarity between these T-cell lineages in future studies.

Our present results firmly establish the 'thymoid' region of the gill tips as the site of VLRA and VLRC assembly, and indicate that the epithelial microenvironment of this primary lymphoid organ provides the necessary signals to support the development of the two T-cell-like lineages of lymphocytes identified here. The thymoid thus parallels the jawed vertebrate thymus in its support of both TCRαβ<sup>+</sup> and TCRγδ<sup>+</sup> T-cell lineage differentiation. Incomplete intermediate VLRC assemblies were found only in 'thymoid' cells, as shown previously for VLRA assemblies<sup>19</sup>. These findings also indicate that VLRA and VLRC gene assemblies are both associated with CDA1 expression in the 'thymoid' region of the gills, whereas CDA2 expression and VLRB assembly occur in the haematopoietic typhosole and kidney tissues in lamprey larvae<sup>19</sup>. Our analysis of VLRA and VLRC expression and assembly in 'thymoid' cells and in VLRA<sup>+</sup> and VLRC<sup>+</sup> lymphocytes throughout the body suggests that bi-potent lymphoid progenitors undergo a complex pattern of VLRC versus VLRA assembly in which VLRC assembly precedes VLRA assembly in the 'thymoid'.

The findings here reveal functional similarities and differences between the three lymphocyte lineages of lamprey larvae. Unlike VLRB<sup>+</sup> cells, the VLRC<sup>+</sup> and VLRA<sup>+</sup> lymphocytes preferentially populate the pharyngeal and intestinal epithelium. Both thymoid-derived lymphocyte types can be activated by antigen stimulation or by the polyclonal mitogen PHA to undergo lymphoblastoid transformation and cell division, but do not differentiate into VLR-secreting cells, a chief difference to VLRB<sup>+</sup> cells. Notably, the VLRC<sup>+</sup> cells constitute the predominant lymphocyte population in the lamprey epidermis, and the repertoire of VLRC<sup>+</sup> cells derived from skin samples is restricted relative to that for VLRC<sup>+</sup> cells in other tissue sites. These interesting findings, which bring to mind the canonical γδ TCR expressed by dendritic epithelial T cells in mice<sup>10</sup> and the antigen-specific αβ T cells that patrol the skin of mice after herpes virus infection<sup>20</sup>, emphasize the need to elucidate the mode of antigen recognition by the VLRC- and VLRA-bearing cells.

In conjunction with the previously documented dichotomy of B- and T-cell-like lymphocytes in lamprey, these results offer fresh insight into the evolution of the alternative adaptive immune systems of jawless and jawed vertebrates. Representative species of the two vertebrate lineages use different sets of genes and different combinatorial mechanisms to construct diverse antigen recognition receptors, but they show the same basic principle of lymphocyte differentiation along two distinct T-cell-like lineages and one B-cell-like lineage. These composite findings suggest that the basic genetic program regulating the developmental pathways of the three lymphocyte lineages must have already existed in a common ancestor for all vertebrates (Supplementary Fig. 6) and this constitutes a fundamental organizing principle for lymphocyte-based adaptive immune systems.

## METHODS SUMMARY

Two outbred species of lamprey larvae (*Petromyzon marinus* and *Lampetra plaineri*) were captured in the wild and housed in sand-lined aquaria. Methods were

as described previously<sup>5,19</sup>, except as noted below. VLRC complementary DNA constructs<sup>9</sup> were cloned into a fusion vector to produce chimaeric proteins of the VLRC antigen-binding domain and stalk region fused to the constant region of human immunoglobulin-G1 (IgG1). Lymphocytes from mice immunized with the chimaeric proteins were used for hybridoma production, and supernatants were screened for antibody reactivity against recombinant VLRC. Monoclonal antibodies specific for VLRC were used together with anti-VLRA and -VLRB antibodies for characterizing and purifying VLR-expressing cells. Target gene sequences derived from the lamprey genome database<sup>21</sup> were used for quantitative real-time PCR analysis of genes and transcripts of the purified lymphocyte populations.

**Full Methods** and any associated references are available in the online version of the paper.

**Received 22 March; accepted 15 July 2013.**

**Published online 11 August 2013.**

1. Flajnik, M. F. & Kasahara, M. Origin and evolution of the adaptive immune system: genetic events and selective pressures. *Nature Rev. Genet.* **11**, 47–59 (2010).
2. Litman, G. W., Rast, J. P. & Fugmann, S. D. The origins of vertebrate adaptive immunity. *Nature Rev. Immunol.* **10**, 543–553 (2010).
3. Tonegawa, S. Somatic generation of antibody diversity. *Nature* **302**, 575–581 (1983).
4. Davis, M. M. & Bjorkman, P. J. T-cell antigen receptor genes and T-cell recognition. *Nature* **334**, 395–402 (1988).
5. Guo, P. *et al.* Dual nature of the adaptive immune system in lampreys. *Nature* **459**, 796–801 (2009).
6. Hirano, M., Das, S., Guo, P. & Cooper, M. D. The evolution of adaptive immunity in vertebrates. *Adv. Immunol.* **109**, 125–157 (2011).
7. Boehm, T. *et al.* VLR-based adaptive immunity. *Annu. Rev. Immunol.* **30**, 203–220 (2012).
8. Kasamatsu, J. *et al.* Identification of a third variable lymphocyte receptor in the lamprey. *Proc. Natl Acad. Sci. USA* **107**, 14304–14308 (2010).
9. Das, S. *et al.* Organization of lamprey variable lymphocyte receptor C locus and repertoire development. *Proc. Natl Acad. Sci. USA* **110**, 6043–6048 (2013).
10. Vantourout, P. & Hayday, A. Six-of-the-best: unique contributions of gammadelta T cells to immunology. *Nature Rev. Immunol.* **13**, 88–100 (2013).
11. Alder, M. N. *et al.* Antibody responses of variable lymphocyte receptors in the lamprey. *Nature Immunol.* **9**, 319–327 (2008).
12. Herrin, B. R. *et al.* Structure and specificity of lamprey monoclonal antibodies. *Proc. Natl Acad. Sci. USA* **105**, 2040–2045 (2008).
13. Melichar, H. J. *et al.* Regulation of γδ versus αβ T lymphocyte differentiation by the transcription factor SOX13. *Science* **315**, 230–233 (2007).
14. Rothenberg, E. V., Moore, J. E. & Yui, M. A. Launching the T-cell-lineage developmental programme. *Nature Rev. Immunol.* **8**, 9–21 (2008).
15. Nakajima, S., Roswit, W. T., Look, D. C. & Holtzman, M. J. A hierarchy for integrin expression and adhesiveness among T cell subsets that is linked to TCR gene usage and emphasizes V delta 1+ gamma delta T cell adherence and tissue retention. *J. Immunol.* **155**, 1117–1131 (1995).
16. Wesch, D. *et al.* Direct costimulatory effect of TLR3 ligand poly(I:C) on human γδ T lymphocytes. *J. Immunol.* **176**, 1348–1354 (2006).
17. Wilson, K. C., Center, D. M. & Cruikshank, W. W. The effect of interleukin-16 and its precursor on T lymphocyte activation and growth. *Growth Factors* **22**, 97–104 (2004).
18. Wesch, D., Peters, C., Oberg, H. H., Pietschmann, K. & Kabelitz, D. Modulation of gammadelta T cell responses by TLR ligands. *Cell. Mol. Life Sci.* **68**, 2357–2370 (2011).
19. Bajoghli, B. *et al.* A thymus candidate in lampreys. *Nature* **470**, 90–94 (2011).
20. Ariotti, S. *et al.* Tissue-resident memory CD8<sup>+</sup> T cells continuously patrol skin epithelia to quickly recognize local antigen. *Proc. Natl Acad. Sci. USA* **109**, 19739–19744 (2012).
21. Smith, J. J. *et al.* Sequencing of the sea lamprey (*Petromyzon marinus*) genome provides insights into vertebrate evolution. *Nature Genet.* **45**, 415–421 (2013).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank C. L. Turnbough Jr for providing *B. anthracis* exosporium, H. Yi for help with electron microscopy, S. A. Durham and R. E. Karaffa II for help with cell sorting, S. Holland for help with gene orthology analysis, Q. Han for help with cloning, and B. R. Herrin, M. Kasahara and Y. Sutoh for suggestions and discussion. M.H., P.G., N.M., S.D. and M.D.C. are supported by National Institutes of Health grants (R01AI072435 and R01GM100151) and the Georgia Research Alliance; M.S. and T.B. are supported by the Max Planck Society.

**Author Contributions** M.H., P.G., N.M., M.S., S.D., T.B. and M.D.C. designed the research, analysed data and wrote the paper; M.H., P.G., N.M., M.S., S.D. and T.B. carried out the research.

**Author Information** Sequence data have been deposited in GenBank/EMBL/DBJ databases under accession numbers KF385949–KF385955. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.D.C. (max.cooper@emory.edu).

## METHODS

**Animals, antigens and mitogens.** Two outbred species of lamprey larvae (*Petromyzon marinus* and *Lampetra planeri*) were captured in the wild and housed in sand-lined aquaria. *Bacillus anthracis* exosporium was provided by C. L. Turnbough Jr. Lampreys were immunized with 10 µg of *B. anthracis* exosporium. Animals were injected with 25 µg poly(I:C) (Sigma) or 25 µg phytohaemagglutinin (PHA-L, Sigma).

**Production of anti-VLRC monoclonal antibodies.** VLRC complementary DNA constructs<sup>9</sup> were cloned into a fusion vector to produce chimaeric proteins of the VLRC antigen-binding domain and stalk region fused to the constant region of human immunoglobulin-G1 (IgG1). Lymphocytes from mice immunized with the chimaeric proteins were used for hybridoma production, and supernatants were screened for antibody reactivity against recombinant VLRC. Monoclonal antibodies specific for VLRC were used together with anti-VLRA and -VLRB antibodies for characterizing and purifying VLR-expressing cells.

**ELISA.** ELISA (enzyme-linked immunosorbent assay) plates were coated with VLRA-IgG1-Fc<sup>5</sup> or VLRC-IgG1-Fc fusion proteins for 12 h at 4 °C and blocked with 1% BSA in PBS for 3 h at 37 °C. Hybridoma culture supernatants were added for 2 h at 37 °C. VLR reactivity of antibodies was detected using alkaline phosphatase-conjugated goat anti-mouse immunoglobulin antibodies (Southern Biotech), developed with phosphatase substrate (Sigma) and read at 405 nm (Versamax microplate reader, Molecular Devices).

**Flow cytometric analysis and sorting of VLRA, VLRB and VLRC lymphocytes.** Leukocyte isolation from blood and tissues and staining for flow cytometry proceeded as described<sup>5</sup>. In brief, leukocytes from blood and tissues were stained with anti-VLRA rabbit polyclonal serum (R110), anti-VLRB mouse monoclonal antibody (4C4) and biotinylated anti-VLRC mouse monoclonal antibodies (1B4, 3A5, 10A5, 11B5) and matched secondary reagents. Flow cytometric analysis was performed on a CyAn ADP (Dako) or Accuri C6 (BD Biosciences) flow cytometers. VLRA<sup>+</sup>, VLRB<sup>+</sup>, VLRC<sup>+</sup> and VLR triple-negative cells in the lymphocyte gate were sorted on BD FACS Aria II (BD Bioscience) for genomic and quantitative RT-PCR analysis. The purity of the sorted cells was 96.5 ± 1.5% (VLRA<sup>+</sup>), 96.1 ± 1.4% (VLRB<sup>+</sup>), 98.7 ± 1.3% (VLRC<sup>+</sup>) and 97.9 ± 1.4% (triple-negative).

**Immunofluorescence microscopy.** Tissue samples were prepared and stained as described previously<sup>19</sup> using anti-VLRA rabbit serum (R110) and anti-VLRC monoclonal antibodies (3A5). Fluorescence microscopy was performed with a Zeiss Axiovert 200M and images were processed with Adobe Photoshop (Adobe Systems). For cell-number counts from intestinal epithelium and epidermis, the total numbers of positive cells from single 7-µm sections were counted for both tissues in a total of five larvae.

**Proliferation assay.** Lampreys stimulated with antigen, mitogen or poly(I:C) were injected with 5 µg of 5-ethynyl-2'-deoxyuridine (Invitrogen) in 40 µl 0.67 × PBS and returned to aquaculture for 24 h before collecting leukocytes for flow cytometry analysis as described<sup>5</sup>.

**Lymphocyte counts.** Lymphocytes were prepared from lamprey blood and kidneys by passage through 70 µm cell strainers (BD Biosciences), producing single-cell suspensions. Cells were washed and resuspended in 1 ml 0.67 × PBS. Total cells were counted in 25 µl of each sample on an Accuri C6 flow cytometer. The total number of cells in the 'lymphocyte gate' was calculated by the formula: total lymphocyte number = number of cells in the lymphocyte gate × 40 (dilution factor). Lymphocytes were then stained with anti-VLRA, -VLRB and -VLRC antibodies to determine the percentages of VLRA<sup>+</sup>, VLRB<sup>+</sup> and VLRC<sup>+</sup> cells, which were used to calculate the numbers of lymphocytes of each type.

**VLR expression in transfected HEK-293T cells.** VLRB- and VLRC-pIRESpuo2 plasmids were transfected into HEK-293T cells cultured in DMEM containing 5% FBS using linear polyethylenimine (PEI), MW 25,000 (Polysciences) at a 3:1 PEI:DNA ratio. Cells were separated from supernatants 48 h after transfection by centrifugation at 300g and lysed in 1% NP-40 lysis buffer.

**Immunoblot analysis.** Samples were separated on 11% non-reducing SDS-PAGE gels before transfer to polyvinylidene fluoride membranes (Millipore). Membranes were blocked overnight with 5% skimmed milk (US Biological) and incubated with anti-VLRB (4C4) or anti-VLRC (3A5) antibodies for 1 h. After five washes with 0.5% Tween-20 in 1 × PBS, membranes were incubated with horseradish peroxidase (HRP)-conjugated goat anti-mouse immunoglobulin polyclonal antibodies (Dako) and washed. Blots were developed using SuperSignal West Pico Chemiluminescent substrate (Thermo Scientific).

**Genomic PCR.** Genomic DNA was extracted from VLRA<sup>+</sup>, VLRB<sup>+</sup>, VLRC<sup>+</sup> and VLR triple-negative-sorted lymphocytes using DNeasy kit (QIAGEN). Genomic PCR was carried out using primers VLRA-F and VLRA-R, VLRB-F and VLRB-R or VLRC-F and VLRC-R (Expand Long Template, Roche). Primers are listed in Supplementary Table 1.

**RT-PCR.** Total RNA was extracted from sorted cells of lamprey blood and tissues (kidneys, typhlosole, gill and skin) and reverse-transcribed using random hexamers (Invitrogen). KOD Hot Start DNA Polymerase (TOYOBO) and Ex Taq were used to amplify the VLR genes. When necessary, the PCR products were cloned into pBluescript (Stratagene) or pCR4-TOPO (Invitrogen) vector and sequenced. Primers used for PCR are listed in Supplementary Table 1.

**VLRC assemblies from *Lampetra planeri*.** Genomic DNA was procured from the thymoid region of gill filaments of *Lampetra planeri* larvae by laser-capture micro-dissection, essentially as described<sup>19</sup>. VLRC genes were amplified using high-fidelity Phusion Taq polymerase (Invitrogen) and primers VLRC5.1 and VLRC3 (Supplementary Table1), cloned into the pGEM-T vector (Promega) and sequenced; in some cases, sequences were cloned after a second amplification step using primers VLRC5.2 and VLRC3. Non-functional sequences exhibited either internal stop codons, frame-shift mutations, or consisted of partial assemblies.

**RNA *in situ* hybridization analyses.** RNA *in situ* hybridization analyses were performed on paraffin-embedded tissue sections of *L. planeri* larvae, essentially as described<sup>19</sup>. The *CDA1* probe has been described previously<sup>19</sup>.

**Quantitative real-time PCR.** Target gene sequences were obtained from the lamprey genome database<sup>21</sup> and subsequent *in silico* analysis (including phylogenetic analysis) was performed as described previously<sup>9</sup>. RNA was extracted from each population using RNeasy kits with on-column DNA digestion by DNase I (QIAGEN). First-strand cDNA was synthesized with random hexamer primers and Superscript III (Invitrogen). Quantitative real-time PCR was carried out with SYBR Green on a 7900HT ABI Prism (Applied Biosystems). Quantitative real-time PCR reactions were performed to evaluate the expression of each gene orthologue. Five separate determinations were carried out for heat-map analysis. The value of the target gene expression was normalized to β-actin. The normalized value for each gene was compiled into a heat map (*z*-scores) with three-colour scale. Red, *z* = 1.5; blue, *z* = -1.5; white, *z* = 0 (*z* = (each value - average)/standard deviation). Primers and the values of the target gene are described in Supplementary Tables 1 and 2, respectively.

**Electron microscopy.** Tissue lymphocytes from naive or PHA-stimulated lamprey larvae were sorted on a BD FACS Aria II (BD Biosciences). The VLRC-positive cells were prepared for transmission electron microscopic analysis as described previously<sup>5</sup>.

**Statistical analysis.** Statistical significance was determined by a two-sample Student's *t*-test and Fisher's exact test.

# Cross-neutralization of four paramyxoviruses by a human monoclonal antibody

Davide Corti<sup>1,2</sup>, Siro Bianchi<sup>1</sup>, Fabrizia Vanzetta<sup>1</sup>, Andrea Minola<sup>2</sup>, Laurent Perez<sup>2</sup>, Gloria Agatic<sup>1</sup>, Barbara Guarino<sup>1</sup>, Chiara Silacci<sup>2</sup>, Jessica Marcandalli<sup>2</sup>, Benjamin J. Marsland<sup>3</sup>, Antonio Piralla<sup>4</sup>, Elena Percivalle<sup>4</sup>, Federica Sallusto<sup>2</sup>, Fausto Baldanti<sup>4</sup> & Antonio Lanzavecchia<sup>2,5</sup>

**Broadly neutralizing antibodies reactive against most and even all variants of the same viral species have been described for influenza and HIV-1 (ref. 1). However, whether a neutralizing antibody could have the breadth of range to target different viral species was unknown. Human respiratory syncytial virus (HRSV) and human metapneumovirus (HMPV) are common pathogens that cause severe disease in premature newborns, hospitalized children<sup>2,3</sup> and immune-compromised patients<sup>2,4,5</sup>, and play a role in asthma exacerbations<sup>6</sup>. Although antisera generated against either HRSV or HMPV are not cross-neutralizing<sup>7</sup>, we speculated that, because of the repeated exposure to these viruses, cross-neutralizing antibodies may be selected in some individuals. Here we describe a human monoclonal antibody (MPE8) that potently cross-neutralizes HRSV and HMPV as well as two animal paramyxoviruses: bovine RSV (BRSV) and pneumonia virus of mice (PVM). In its germline configuration, MPE8 is HRSV-specific and its breadth is achieved by somatic mutations in the light chain variable region. MPE8 did not result in the selection of viral escape mutants that evaded antibody targeting and showed potent prophylactic efficacy in animal models of HRSV and HMPV infection, as well as prophylactic and therapeutic efficacy in the more relevant model of lethal PVM infection. The core epitope of MPE8 was mapped on two highly conserved anti-parallel  $\beta$ -strands on the pre-fusion viral F protein, which are rearranged in the post-fusion F protein conformation. Twenty-six out of the thirty HRSV-specific neutralizing antibodies isolated were also found to be specific for the pre-fusion F protein. Taken together, these results indicate that MPE8 might be used for the prophylaxis and therapy of severe HRSV and HMPV infections and identify the pre-fusion F protein as a candidate HRSV vaccine.**

HRSV and HMPV are human paramyxoviruses that differ in their genomic organization and that share only 33% amino acid identity in the F protein, the primary target of neutralizing antibodies<sup>8,9</sup>. Polyclonal antisera (RespiGam) were used to prevent HRSV infection in premature newborns or children at risk and were substituted in 1997 by the anti-F humanized monoclonal antibody palivizumab (Synagis), which is currently the only anti-infective monoclonal antibody approved for human use<sup>10,11</sup>. However, palivizumab is not effective therapeutically<sup>12</sup> and has been shown to select escape mutants *in vivo*<sup>13</sup>. The limited efficacy of palivizumab, the lack of a treatment for HMPV and the lack of effective vaccines highlight the medical need for new therapies against these prominent human pathogens.

To search for a HRSV and HMPV cross-reactive neutralizing antibody, we screened a cohort of 200 blood donors and selected 7 donors with high serum neutralizing antibody titres against both viruses. A total of 114,000 IgG<sup>+</sup> memory B cells were isolated and immortalized using Epstein-Barr Virus (EBV) and CpG, as previously described<sup>14</sup>. The B cell supernatants were analysed for their capacity to neutralize infection by HRSV A/A2/61 and subsequently rescreened

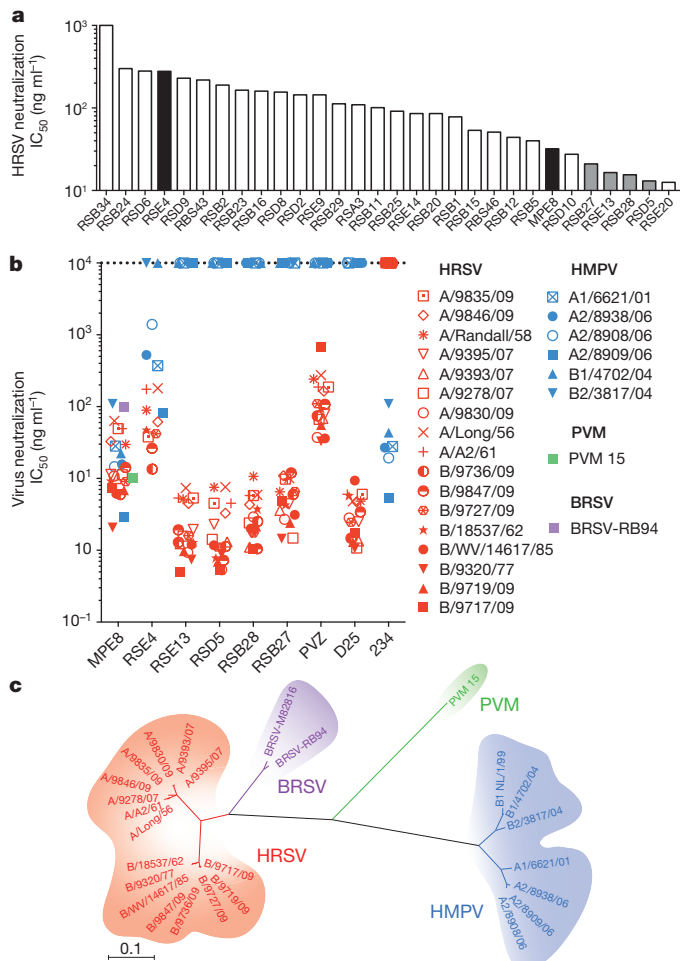
on HMPV A1/6621/01. Thirty HRSV-neutralizing antibodies were isolated and two of them, named MPE8 and RSE4, were found to cross-neutralize HMPV infection (Fig. 1a and Supplementary Table 1). MPE8, RSE4, as well as four HRSV-specific neutralizing antibodies (RSE13, RSD5, RSB27 and RSB28) selected for their high potency, were compared to palivizumab and two previously described antibodies, D25 (specific for HRSV)<sup>15</sup> and 234 (specific for HMPV)<sup>16</sup>, for their capacity to neutralize a panel of HRSV and HMPV strains, as well as the two related animal paramyxoviruses PVM and BRSV (Fig. 1b, c). Strikingly, MPE8 neutralized many HRSV and HMPV isolates representative of all subgroups, as well as PVM and BRSV, with IC<sub>50</sub> (half-maximum inhibitory concentration) values ranging from 2 to 110 ng ml<sup>-1</sup> (mean 23 ng ml<sup>-1</sup>). When compared to palivizumab, MPE8 was on average eightfold more potent in neutralizing HRSV (Fig. 1b and Supplementary Table 2). RSE4 showed a lower potency and breadth, as it failed to neutralize HMPV B strain viruses. The four HRSV-specific antibodies neutralized HRSV with IC<sub>50</sub> values 27- to 132-fold lower than palivizumab. The neutralizing antibodies were also tested for their capacity to prevent cell-to-cell viral spread, which has been reported to be a distinct property of anti-HRSV antibodies, independent of the neutralizing activity<sup>17</sup>. In this assay MPE8 showed a 50% spreading inhibiting capacity (IS<sub>50</sub>) comparable to that of the most potent neutralizing antibodies RSD5, RSE13, RSB27 and RSB28 (Supplementary Figs 1 and 2). Taken together, these results indicate that MPE8 can potently neutralize four different paramyxoviruses and its mechanism of action involves both viral neutralization as well as inhibition of viral spread and syncytia formation.

MPE8 utilizes VH3-31\*01 and VL1-40\*01, has a 17 amino acid heavy chain complementarity determining region 3 (HCDR3), and carries 9 amino acid substitutions in VH and 4 in VL (Fig. 2a). To analyse the role of somatic mutations, we produced MPE8 variants that were completely or partially reverted to the germline sequence. The fully germlined version of MPE8 neutralized HRSV A and B strains, but failed to neutralize HMPV (Fig. 2b). HMPV binding required mutations in both heavy and light chains. In particular, the reversion of the somatic mutation at position 56 (D56G) in light chain complementarity determining region 2 (LCDR2) was sufficient to abolish neutralization of all HMPV strains, whereas reversion of the somatic mutation at position 65 in LCDR2 (N65S) and 109 in LCDR3 (R109S) had minor effects. Finally, reversion of the somatic mutation at position 113 (N113S), which introduced a glycosylation site in LCDR3, did not affect antibody breadth and potency (Fig. 2b and Supplementary Fig. 3). These findings indicate that the naive B cell that generated MPE8 was specific for HRSV and that the cross-reactivity to HMPV was acquired through somatic mutations, in particular a G to D mutation at position 56 in LCDR2.

All the 30 HRSV neutralizing antibodies isolated selectively stained 293F cells expressing the F, but not the G, protein of HRSV

<sup>1</sup>Humabs BioMed SA, Via Mirasole 1, 6500 Bellinzona, Switzerland. <sup>2</sup>Institute for Research in Biomedicine, Università della Svizzera Italiana, Via Vincenzo Vela 6, 6500 Bellinzona, Switzerland. <sup>3</sup>University of Lausanne, Service de Pneumologie, CHUV, Rue du Bugnon 46, 1011 Lausanne, Switzerland. <sup>4</sup>Policlinico San Matteo, IRCCS, Piazzale Camillo Golgi 1, 27100 Pavia, Italy. <sup>5</sup>Institute of Microbiology, ETH Zurich, Wolfgang-Pauli-Strasse 10, 8093 Zurich, Switzerland.

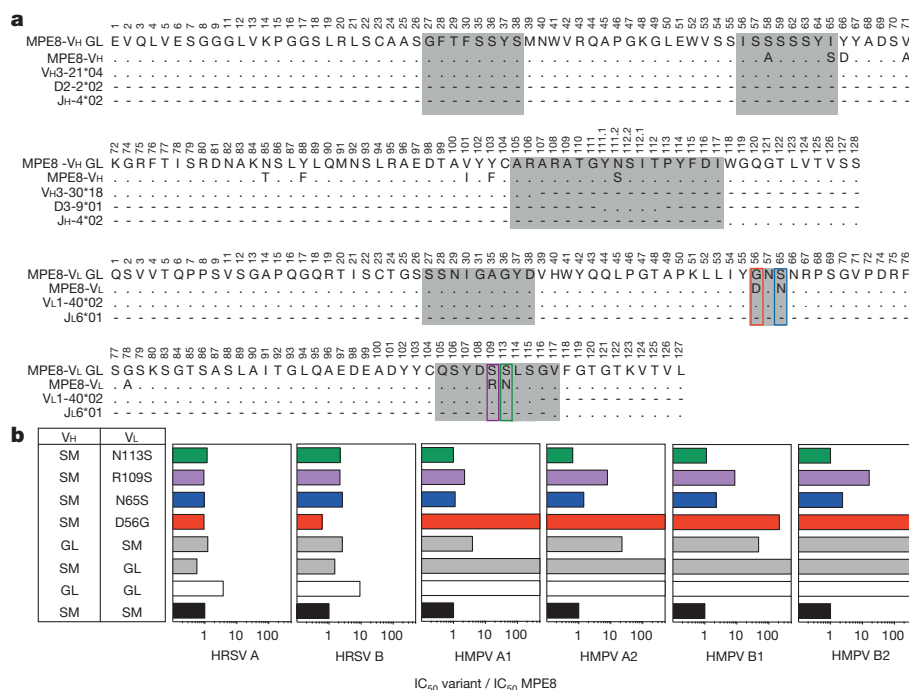




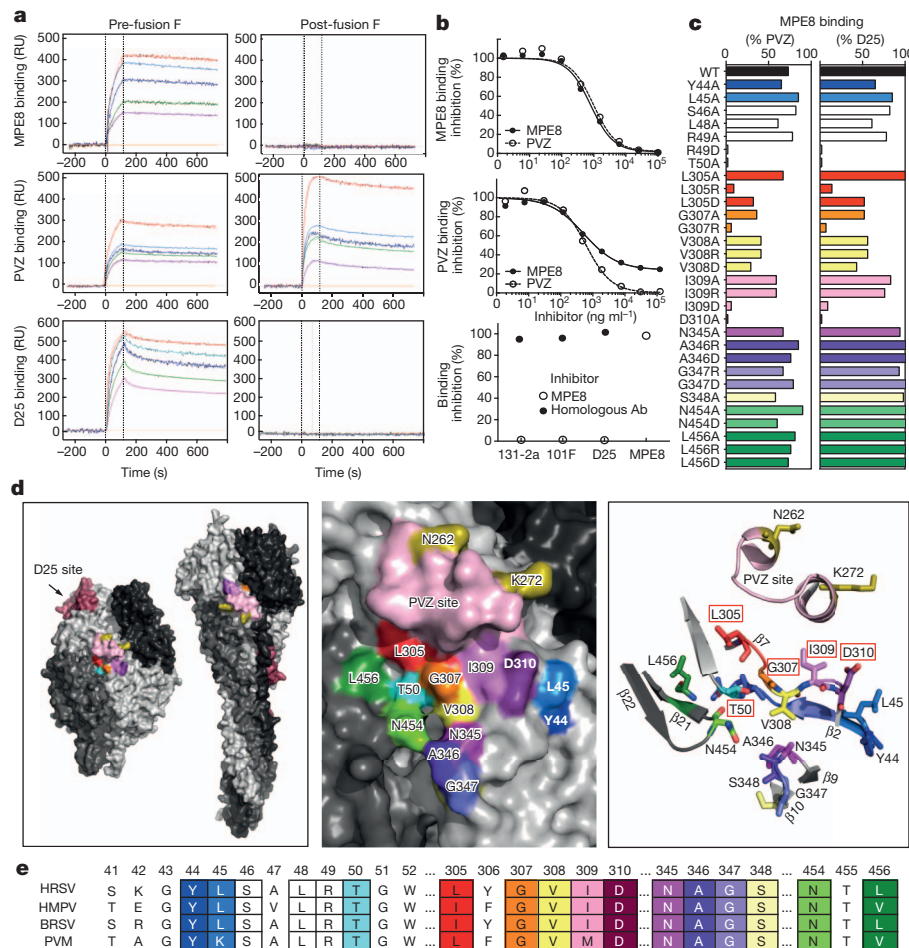
(Supplementary Table 1), indicating that the F protein is the major target of human neutralizing antibodies. Because the HRSV pre-fusion F protein is very unstable, we engineered a stabilized soluble form by: 1)

**Figure 1 | MPE8 cross-neutralizes four human and animal paramyxoviruses.** **a**, Thirty human HRSV neutralizing monoclonal antibodies were isolated from memory B cells. The histogram shows the antibodies ranked according to their  $IC_{50}$  values against the HRSV A/A2/61 strain. The six antibodies further characterized are highlighted in grey (HRSV specific) or in black (HRSV and HMPV specific). **b**, Neutralization of HRSV (red), HMPV (blue), PVM (green) and BRSV (purple) by the six selected antibodies and by palivizumab (PVZ), D25 and 234 as controls.  $IC_{50}$  values ( $ng\ ml^{-1}$ ) were determined by immunofluorescence analysis and measured by nonlinear fitting.  $IC_{50}$  values above  $10^4\ ng\ ml^{-1}$  were scored as negative. The data with complete viral strain designations and accession numbers are listed in Supplementary Table 2. **c**, Dendrogram showing the protein distance of the HRSV, HMPV, PVM and BRSV F proteins tested. The amino acid distance scale is indicated with a value of 10% distance. Experiments were performed in quadruplicates. Data are representative of at least two independent experiments. RSE20 was not further characterized because it failed to neutralize HRSV B strains.

introduction of inter-monomeric disulphide bridges<sup>18</sup>, 2) deletion of the fusion peptide region, and 3) substitution of the transmembrane region with a T4 fibrin motif ('foldon') to stabilize the trimeric structure (Supplementary Figs 4 and 5). We also produced the post-fusion F protein, which has been developed as a subunit HRSV vaccine<sup>19</sup>. Electron microscopy of HRSV post-fusion F protein showed a field of uniform 'crutch'-shaped molecule and rosette aggregates characteristic of post-fusion F proteins<sup>20</sup>. Conversely, HRSV pre-fusion F protein showed a homogeneous arrangement of 'cone'-shaped molecules, characteristic of F proteins of other paramyxoviruses in the pre-fusion state<sup>21</sup> (Supplementary Fig. 6). When tested by surface plasmon resonance (SPR), palivizumab bound to both pre- and post-fusion proteins with similar kinetics (Fig. 3a). In contrast, 26 out of the 30 neutralizing antibodies isolated (including MPE8, RSD5, RSE13, RSB27, RSB28 and D25), selectively bound to the pre-fusion F protein (Fig. 3a and Supplementary Table 1). Using cross-competition SPR experiments and the panel of available antibodies, we identified seven antigenic sites (S1 to S7) on the pre-fusion F protein, four of which were targeted by the potent neutralizing antibodies (Supplementary Fig. 7). MPE8 and palivizumab cross-competed for binding to the S1 site of the pre-fusion F protein, indicating that their cognate epitopes are in close proximity on the pre-fusion F protein, however, the epitope recognized by MPE8 is not accessible on the post-fusion F protein. This conclusion is consistent



**Figure 2 | Somatic mutations broadening the MPE8 specificity.** **a**, Alignment of  $V_H$  and  $V_L$  amino acid sequences of MPE8 with its germline configuration. Shown are V, D and J sequences determined using the IMGT database (<http://www.imgt.org/>). The CDR regions are highlighted in grey. Dots indicate identical residues; dashes indicate that residues are absent. **b**, Neutralization of HRSV and HMPV A and B strains by different versions of MPE8 in which all or individual somatic mutations in the  $V_H$  or  $V_L$  chain were reverted to the germline. Shown is the ratio between the  $IC_{50}$  values of the variants and  $IC_{50}$  values of MPE8. The mutations indicated represent reversions to the corresponding germline residues. GL, germlined  $V_H$  or  $V_L$ ; SM, somatically mutated  $V_H$  or  $V_L$ . The MPE8 (N113S) mutant was used in all further experiments. Coloured boxes in panel a reflect the colours used in the panel b bar charts.



**Figure 3 | MPE8 binds to a conserved site on the pre-fusion F protein.**

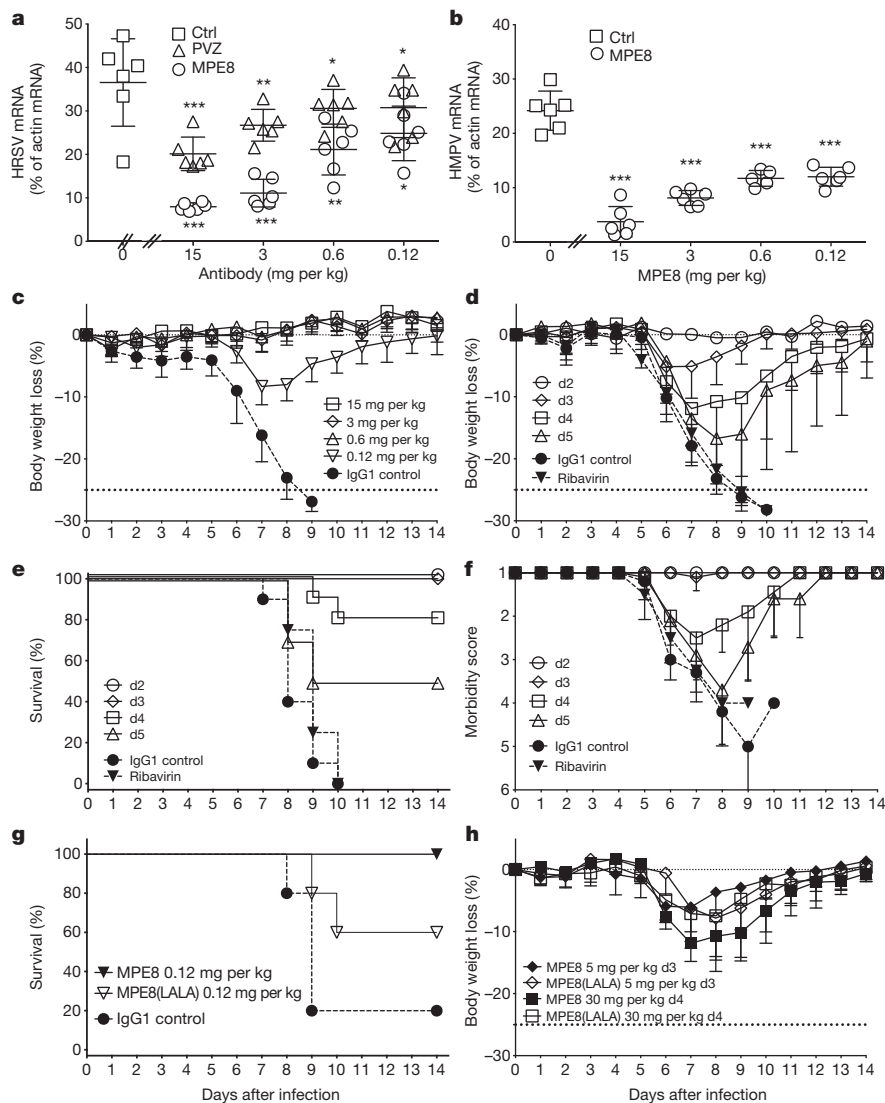
**a**, Binding of MPE8, palivizumab (PVZ) and D25 to pre-fusion and post-fusion HRSV F proteins as measured by SPR. RU, resonance units. **b**, Cross-competition for binding to HRSV-infected cells. MPE8 binding is completely inhibited by PVZ (upper panel), whereas PVZ binding is only partially inhibited by MPE8 (middle panel); binding of antibodies 131-2a, 101F and D25 is not inhibited by MPE8. **c**, Binding of MPE8 to cells transfected with HRSV F genes carrying specific mutations. The results are expressed as percentage binding of PVZ and D25. The colour codes reflect the position of the mutated residues as shown in panel **d**. **d**, Left panel, model of the trimeric pre-fusion

(Protein Data Bank (PDB) code 4JHW) and post-fusion (PDB code 3RKI) HRSV F proteins showing the location of the D25 epitope (brown), the PVZ site (light pink) and the MPE8 site (as defined by residues T50, L305, G307, I309 and D310). The three monomers are depicted with different shades of grey. Middle and right panels show a close up of the MPE8 and PVZ sites on the pre-fusion F protein. The critical residues for MPE8 binding identified by mutagenesis are boxed in red. Shown in gold are the two residues N262 and K272 which are found mutated in the PVZ MARMs. **e**, Alignment of HRSV, HMPV, BRSV and PVM F protein residues which have been mutated.

with the finding that palivizumab completely inhibited binding of MPE8 to HRSV-infected cells (which display both pre- and post-fusion F proteins), whereas MPE8 only partially inhibited binding of palivizumab (Fig. 3b). In addition, MPE8 did not cross-compete with three antibodies that were mapped to distinct sites on the F protein<sup>22,23</sup> (Fig. 3b).

MPE8 reacted in western blots with HRSV and HMPV F proteins under non-reducing conditions (Supplementary Fig. 8), suggesting that the epitope is at least transiently accessible in the presence of anionic detergents, allowing the antibody to 'lock-on'. We therefore screened a library of 7,095 structured peptides covering the sequence of HRSV F protein and identified a peptide in the amino-terminal region, spanning residues SAVSKGYLSALRTGWYTSVIT, which was bound by MPE8. The core sequence in this peptide (YLSALRTGW) identifies, on the surface of the recently described pre-fusion F protein structure<sup>23</sup>, a patch of residues that are conserved in both HRSV and HMPV F proteins (Supplementary Fig. 9). To refine the MPE8 epitope mapping, we expressed in 293F cells HRSV F protein mutants located in the anti-parallel  $\beta$ 2 and  $\beta$ 7 strands (residues 44–50 and 305–310) and in two neighbouring loops, corresponding to residues 345–348 of the same monomer and 454–456 of the adjacent monomer (Fig. 3c–e). Complete loss of binding by MPE8 was observed with two alanine

mutants of the surface-exposed residues T50 and D310, as well as with the non-conservative aspartate substitution at R49, which lies below T50. Substantial loss of binding was also caused by mutations in surface residues L305R, G307R and I309D. Interestingly, the  $\beta$ -strands are reoriented in the post-fusion F protein structure and residues T50 and L305 are not surface exposed. All mutants tested were bound by a panel of antibodies, including palivizumab, D25, RSD5 and RSE13, indicating that the mutations were not disruptive to other epitopes. As a second approach for epitope mapping, we attempted to isolate monoclonal antibody resistant mutants (MARMs). HRSV MARMs were isolated at high frequency (1/200,000 TCID<sub>50</sub> (50% tissue culture infective dose)) in the presence of palivizumab and carried mutations at positions 262 and 272, consistent with previous findings<sup>24,25</sup>. In contrast, neither HRSV nor HMPV MARMs were isolated in the presence of MPE8. Furthermore, all palivizumab MARMs were neutralized by MPE8, consistent with the two antibodies recognizing close, yet distinct, epitopes in the S1 site of the pre-fusion F protein (Supplementary Fig. 10 and Fig. 3d). Taken together, the above results indicate that MPE8 binds to a highly conserved epitope that is displayed on the surface of the pre-fusion, but not of the post-fusion, HRSV and HMPV F proteins. This epitope is close to, but distinct from, that recognized by



**Figure 4 | Prophylactic and therapeutic efficacy of MPE8 against HRSV, HMPV and PVM.**

**a, b,** HRSV and HMPV prophylaxis. BALB/c mice (*n* = 6) were injected with different doses of MPE8 (circles), palivizumab (PVZ, triangles) or the diluent as a control (Ctrl, squares) four hours before intranasal (i.n.) infection with  $10^6$  p.f.u. of HRSV A/A2/61 (**a**) or HMPV A/6621/04 (**b**). Mice were euthanized on day 5 and viral messenger RNA levels were determined by quantitative PCR (qPCR). Two-tailed *P* value was calculated with the unpaired Student's *t*-test. \*\*\**P* < 0.0001; \*\**P* = 0.001 to 0.01; \**P* = 0.01 to 0.05. **c,** PVM prophylaxis. BALB/c mice (*n* = 5) were injected with different doses of MPE8 or control HJ16 anti-HIV antibody (Ctrl) and challenged (i.n.) after 4 h with  $3 \times \text{MLD}_{50}$  (median lethal dose at 50%) of PVM 15. Shown is the percentage of body weight loss. **d–f,** PVM therapy. BALB/c mice (*n* = 10) were infected with PVM as above and injected with MPE8 at 30 mg per kg on day 2, 3, 4 or 5 after infection. Control mice were treated on day 0 with HJ16. A group of mice was treated with ribavirin intraperitoneal (i.p.) administration at 37.5 mg per kg, twice a day, from day 3 onwards. Shown is the percentage of body weight loss (**d**), per cent survival (**e**) and morbidity score (**f**). **g, h,** Role of antibody Fc in PVM prophylaxis and therapy. Comparison of MPE8 and MPE8(LALA) tested in a prophylactic setting at 0.12 mg per kg (**g**) or in a therapeutic setting at 5 mg per kg on day 3 or 30 mg per kg on day 4 after infection (**h**) (*n* = 5). All error bars indicate s.d.

palivizumab, is centred around the  $\beta 2/\beta 7$  antiparallel strands and relies on residues T50, D305, G307, I309 and D310, which are highly conserved amongst 551 virus strains, comprising 364 HRSV, 162 HMPV, 8 BRSV and 5 PVM strains (Supplementary Fig. 11).

Finally, we set out to test the *in vivo* efficacy of MPE8 in prophylactic and therapeutic settings. In the HRSV mouse model, MPE8 was five- to tenfold more potent than palivizumab in reducing HRSV lung titres and was effective at concentrations as low as 0.12 mg per kg of body weight. Similar results were obtained in the HMPV mouse model (Fig. 4a, b). However, when tested at 15 mg per kg in a therapeutic setting in *Stat1*-deficient mice, which sustain higher HRSV replication<sup>26</sup>, MPE8, RSE13 and palivizumab showed comparable activity in being effective only when administered on day 1 (Supplementary Fig. 12). These findings highlight the well-known limitations of the mouse model of HRSV infection, which requires a high level of virus challenge to achieve only a low and transient virus replication. We therefore tested MPE8 in the more relevant PVM mouse model, in which a small inoculum causes acute lower respiratory tract infection and lethality, recapitulating the features of severe HRSV and HMPV infection in humans<sup>27</sup>. In a prophylactic setting, MPE8 fully protected mice from PVM lethality at 0.12 mg per kg and from body weight loss at 0.6 mg per kg (Fig. 4c). Furthermore, in a therapeutic setting MPE8 completely protected from lethality when administered up to 3 days after infection both at 30 and 5 mg per kg and conferred significant protection when given on day 4 or

5 at 30 mg per kg (Fig. 4d–f). In this system, ribavirin, which is the only approved standard of care for therapy in humans<sup>28</sup>, was ineffective, as previously described<sup>29</sup>.

To address the role of effector mechanisms *in vivo*, we compared the IgG1 MPE8 with a mutant (L234A, L235A, referred to here as LALA) that lacks complement and Fc receptor binding (MPE8(LALA)). The two variants showed equivalent *in vitro* neutralizing activity (Supplementary Fig. 13). When administered in limiting amounts in a prophylactic setting (0.12 mg per kg), MPE8(LALA) showed a severely reduced efficacy, consistent with an important contribution of effector mechanisms *in vivo* (Fig. 4g). In contrast, when used in a therapeutic setting, MPE8(LALA) mutant was as effective as MPE8 at all doses tested (Fig. 4h). These findings suggest that the therapeutic efficacy of MPE8 in the PVM model may be due to a combination of factors, such as the potent neutralizing and spreading inhibition activity, the failure to select escape mutants, and the selective recognition of the pre-fusion F protein, which avoids the consumption of the antibody by the abundant post-fusion F protein acting as a decoy. Surprisingly, the therapeutic efficacy of MPE8 does not require effector functions, suggesting that the therapeutic activity *in vivo* relies primarily on viral neutralization and inhibition of viral spread. This is supported by the finding that MPE8 administered therapeutically on days 1, 2 and 3 after infection can efficiently block further increase in lung viral RNA (Supplementary Fig. 14).



In conclusion, MPE8 is the first example of an antibody that potently cross-neutralizes up to four different viral species and shows prophylactic and therapeutic efficacy *in vivo*. MPE8 appears to have been selected in a stepwise fashion by HRSV infection followed by HMPV infection in a donor who had high titres of neutralizing antibodies against both viruses. According to the original antigenic sin theory, the naive MPE8 B cell clone was initially primed by HRSV and generated a variant (carrying the G56D mutation), which was cross-reactive with HMPV and was therefore recalled by HMPV.

Given the large prevalence and overlapping winter epidemics of HRSV and HMPV infections, as well as the possibility of co-infections<sup>30</sup>, there is need to develop an effective prophylactic and therapeutic approach against both viruses. Owing to its high potency and breadth, MPE8 could be developed for prophylaxis of both HRSV and HMPV infections not only in newborns, but also in immune-compromised adults. Furthermore, it would be important to assess the potential use of this antibody in a therapeutic setting, for instance to prevent the spreading of HRSV or HMPV from the upper to the lower respiratory tract in high-risk individuals, such as haematopoietic stem-cell transplant patients, or to treat patients with severe lower respiratory tract infections. Finally, the finding that most of the human neutralizing antibodies selectively recognize the pre-fusion F protein provides relevant insights for vaccine design. Our results suggest that a vaccine based on a stabilized pre-fusion F protein might be able to selectively elicit neutralizing antibodies, including MPE8-like antibodies, whose breadth is achievable through a very limited number of somatic mutations.

## METHODS SUMMARY

Peripheral blood CD22<sup>+</sup>IgG<sup>+</sup> B cells were sorted by flow cytometry, plated at 3–5 cells per well and immortalized with EBV and CpG on feeder cells. Culture supernatants of B cells were screened by microneutralization against HRSV A2 virus and the positive cultures were re-screened against HMPV. VH and VL genes from selected cultures were sequenced and evaluated for their potency and breadth against a panel of HRSV, HMPV, PVM and BRSV strains. MPE8 antibody variants in the VH, VL or Fc (LALA mutant) were generated by site-directed mutagenesis or gene synthesis and produced recombinantly in 293F cells. Epitope specificity was determined by scanning peptides, cross-competition studies and staining of 293F cells expressing HRSV F protein mutants. MPE8 prophylactic and therapeutic activity was evaluated in the mouse models of HRSV, HMPV and PVM as described in the Methods. Student's two-tailed paired *t*-test was used for statistical comparisons; *P* values of 0.05 or less were considered significant.

**Full Methods** and any associated references are available in the online version of the paper.

**Received 25 March; accepted 8 July 2013.**

**Published online 18 August 2013.**

1. Corti, D. & Lanzavecchia, A. Broadly neutralizing antiviral antibodies. *Annu. Rev. Immunol.* **31**, 705–742 (2013).
2. Hall, C. B. *et al.* The burden of respiratory syncytial virus infection in young children. *N. Engl. J. Med.* **360**, 588–598 (2009).
3. Edwards, K. M. *et al.* Burden of human metapneumovirus infection in young children. *N. Engl. J. Med.* **368**, 633–643 (2013).
4. Williams, J. V. *et al.* Human metapneumovirus and lower respiratory tract disease in otherwise healthy infants and children. *N. Engl. J. Med.* **350**, 443–450 (2004).
5. Falsey, A. R., Hennessy, P. A., Formica, M. A., Cox, C. & Walsh, E. E. Respiratory syncytial virus infection in elderly and high-risk adults. *N. Engl. J. Med.* **352**, 1749–1759 (2005).
6. Edwards, M. R., Bartlett, N. W., Huxell, T., Openshaw, P. & Johnston, S. L. The microbiology of asthma. *Nature Rev. Immunol.* **10**, 459–471 (2012).
7. Wyde, P. R., Chetty, S. N., Jewell, A. M., Boivin, G. & Piedra, P. A. Comparison of the inhibition of human metapneumovirus and respiratory syncytial virus by ribavirin and immune serum globulin *in vitro*. *Antiviral Res.* **60**, 51–59 (2003).
8. Collins, P. L. & Crowe, J. E. Jr. In *Fields Virology* 6th edn. (eds Knipe D. M. & Howley P. M.) Respiratory syncytial virus and metapneumovirus (Lippincott Williams & Wilkins, 2013).
9. van den Hoogen, B. G. *et al.* A newly discovered human pneumovirus isolated from young children with respiratory tract disease. *Nature Med.* **7**, 719–724 (2001).

10. Groothuis, J. R., Simoes, E. A., Hemming, V. G. & Respiratory Syncytial Virus Immune Globulin Study Group. Respiratory syncytial virus (RSV) infection in preterm infants and the protective effects of RSV immune globulin (RSVIG). *Pediatrics* **95**, 463–467 (1995).
11. The-IMPact-RSV Study Group. Palivizumab, a humanized respiratory syncytial virus monoclonal antibody, reduces hospitalization from respiratory syncytial virus infection in high-risk infants. *Pediatrics* **102**, 531–537 (1998).
12. de Fontbrune, F. S. *et al.* Palivizumab treatment of respiratory syncytial virus infection after allogeneic hematopoietic stem cell transplantation. *Clin. Infect. Dis.* **45**, 1019–1024 (2007).
13. Adams, O. *et al.* Palivizumab-resistant human respiratory syncytial virus infection in infancy. *Clin. Infect. Dis.* **51**, 185–188 (2010).
14. Traggiai, E. *et al.* An efficient method to make human monoclonal antibodies from memory B cells: potent neutralization of SARS coronavirus. *Nature Med.* **10**, 871–875 (2004).
15. Kwakkenbos, M. J. *et al.* Generation of stable monoclonal antibody-producing B cell receptor-positive human memory B cells by genetic programming. *Nature Med.* **16**, 123–128 (2010).
16. Ulbrandt, N. D. *et al.* Isolation and characterization of monoclonal antibodies which neutralize human metapneumovirus *in vitro* and *in vivo*. *J. Virol.* **80**, 7799–7806 (2006).
17. Beeler, J. A. & van Wyke Coelingh, K. Neutralization epitopes of the F glycoprotein of respiratory syncytial virus: effect of mutation upon fusion function. *J. Virol.* **63**, 2941–2950 (1989).
18. Magro, M. *et al.* Neutralizing antibodies against the preactive form of respiratory syncytial virus fusion protein offer unique possibilities for clinical intervention. *Proc. Natl Acad. Sci. USA* **109**, 3089–3094 (2012).
19. Swanson, K. A. *et al.* Structural basis for immunization with postfusion respiratory syncytial virus fusion F glycoprotein (RSV F) to elicit high neutralizing antibody titers. *Proc. Natl Acad. Sci. USA* **108**, 9619–9624 (2011).
20. Calder, L. J. Electron microscopy of the human respiratory syncytial virus fusion protein and complexes that it forms with monoclonal antibodies. *Virology* **271**, 122–131 (2000).
21. Connolly, S. A., Leser, G. P., Yin, H.-S., Jardetzky, T. S. & Lamb, R. A. Refolding of a paramyxovirus F protein from prefusion to postfusion conformations observed by liposome binding and electron microscopy. *Proc. Natl Acad. Sci. USA* **103**, 17903–17908 (2006).
22. McLellan, J. S. *et al.* Structure of a major antigenic site on the respiratory syncytial virus fusion glycoprotein in complex with neutralizing antibody 101F. *J. Virol.* **84**, 12236–12244 (2010).
23. McLellan, J. S. *et al.* Structure of RSV fusion glycoprotein trimer bound to a prefusion-specific neutralizing antibody. *Science* **340**, 1113–1117 (2013).
24. Zhu, Q. *et al.* Natural polymorphisms and resistance-associated mutations in the fusion protein of respiratory syncytial virus (RSV): effects on RSV susceptibility to palivizumab. *J. Infect. Dis.* **205**, 635–638 (2012).
25. Zhao, X. Respiratory syncytial virus escape mutant derived *in vitro* resists palivizumab prophylaxis in cotton rats. *Virology* **318**, 608–612 (2004).
26. Durbin, J. E. *et al.* The role of IFN in respiratory syncytial virus pathogenesis. *J. Immunol.* **168**, 2944–2952 (2002).
27. Domachowske, J. B., Bonville, C. A. & Rosenberg, H. F. Animal models for studying respiratory syncytial virus infection and its long term effects on lung function. *Pediatr. Infect. Dis. J.* **23**, S228–S234 (2004).
28. Shah, J. N. & Chemaly, R. F. Management of RSV infections in adult recipients of hematopoietic stem cell transplantation. *Blood* **117**, 2755–2763 (2011).
29. Bonville, C. A. *et al.* Functional antagonism of chemokine receptor CCR1 reduces mortality in acute pneumovirus infection *in vivo*. *J. Virol.* **78**, 7984–7989 (2004).
30. Greensill, J. *et al.* Human metapneumovirus in severe respiratory syncytial virus bronchiolitis. *Emerg. Infect. Dis.* **9**, 372–375 (2003).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We would like to thank M. Nussenzweig and H. Wardemann for providing reagents for antibody cloning and expression, I. Capua, B. Zecchin and L. Selli for experiments involving BRSV and R. Wepf for electron microscopy. This work was supported by the European Research Council (grant no. 250348 IMMUNExplore) and the Swiss National Science Foundation (grant no. 141254). A.L. is supported by the Helmut Horten Foundation.

**Author Contributions** D.C. designed experiments, analysed the data and wrote the manuscript; S.B., F.V., C.S., B.G. performed and analysed *in vitro* experiments; A.M., G.A., B.J.M. performed and analysed *in vivo* experiments; L.P. and J.M. prepared pre- and post-fusion F proteins and performed SPR experiments; F.B., E.P. and A.P. isolated and sequenced HRSV and HMPV strains; F.S. wrote the manuscript; A.L. provided overall supervision, analysed the data and wrote the manuscript.

**Author Information** The GenBank accession numbers for MPE8 antibody VH and VL are KF314816 and KF314817. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.L. ([lanzavecchia@irb.usi.ch](mailto:lanzavecchia@irb.usi.ch)) or D.C. ([davide.corti@humabs.ch](mailto:davide.corti@humabs.ch)).

## METHODS

**Viruses.** The HRSV viruses used in this study were A/A2/61 (ACO83301), A/Randall/58, A/Long/56 (ACO83302), B/18537/62 (BAA00240), B/WV/14617/85 (ATCC VR-1400), A/9320/77 (AAR14266), A/9846/09 (JX171066), A/9835/09 (JX171067), A/9278/07 (KC618407), A/9393/07 (KC618408), A/9395/07 (KC618409), A/9830/09 (KC618410), B/9719/09 (JX171070), B/9736/09 (JX171068), B/9847/09 (JX171073), B/9717/09 (JX171071), B/9727/09 (JX171069). The HMPV viruses used in this study were A1/6621/01 (KC588902), A2/8938/06 (KC588905), A2/8909/06 (KC588906), A2/8908/06 (KC588907), B1/4702/04 (KC588903) and B2/3817/04 (KC588904). The PVM strain used in this study was obtained from ATCC (catalogue no. VR-25; AAW02840) and expanded by continuous passage in mice. The BRSV strain used in this study was RB93 (BAA00798).

**Sequencing of F proteins from HRSV and HMPV viral isolates.** Viral RNA was extracted from isolates by the automated extraction kit NucliSENS easyMAG (BioMerieux). The HRSV F gene was amplified by the SuperScriptIII One-Step RT-PCR amplification kit (Invitrogen) and specific primers for RSV-A and RSV-B: RSV-A-F1 5'-ATGGAGTTGCCAATCCTCA-3', RSV-A-R1670 5'-CATTGT AAGAAYATGATTARGTRCT-3', RSV-B-F1 5'-ATGGAGYTGCTGATCCAYA GRTYAAGT-3', RSV-B-R1740 5'-CAGGTGGTGTCTTCTCTATTGCTGAAT GC-3'. The RT-PCR conditions for both HRSV types were: incubation at 56 °C for 30 min, denaturation at 95 °C for 2 min, followed by 45 cycles at 94 °C for 30 s, 56 °C for 30 s and 68 °C for 2 min. The HMPV F gene was amplified using specific primers: MPV-F1 5'-ATGTCTTGGAAAGTGRGTATYA-3', HMPV-R1680 5'-TTGCAATGGAGCYTTRCGAGAC-3' and the following conditions: 56 °C for 30 min, denaturation at 95 °C for 2 min, followed by 45 cycles at 94 °C for 30 s, 56 °C for 30 s and 68 °C for 2 min. HRSV purified PCR products were directly sequenced using the following primers: HRSV-A-F1, HRSV-A-F513 5'-ACT ATCCACAAAACAGGCGTAGTCAGC-3', HRSV-A-F1013 5'-ACAGAGGAT GGTACTGTGACAATGCAGG-3', HRSV-A-R1670, HRSV-A-R1229 5'-TGCCA TAGCATGACACAATGGCTCCTA-3', HRSV-B-F1, HRSV-B-F529 5'-GCTGTA GTCAGTCTATCAATGGGGTCC-3', HRSV-B-F1225 5'-TCTCTWGGAGCTA TAGTGTCTATGC-3', HRSV-B-R1740, HRSV-B-R1135 5'-TGTCAGTGTACAA AGGCTGAC-3', HRSV-B-R667 5'-CAATCTGCTGTTCTTCTGCTGGAA-3'. HMPV purified PCR products were sequenced using the following primers: HMPV-F1; HMPV-F574 5'-AGCTTCAGTCAATTCACAGAAG-3'; HMPV-F1075 5'-TACCCATGCAAAGTYAGCACAGG-3', HMPV-R1680, HMPV-R1593 5'-KGTRACACCAATCAGCTCTGGAG-3' and HMPV-R1103 5'-TGTC TTCCTGTGCTRACTTTTC-3'. Cycle sequencing was performed using the BigDye Terminator Cycle-Sequencing kit (Applied Biosystems) in ABI Prism 3100 DNA sequencer (Applied Biosystems).

**Sample collection.** Peripheral blood mononuclear cell (PBMC) and plasma samples were obtained from healthy donors. Memory B cells were isolated from cryopreserved PBMC. All donors gave written informed consent for research use of blood samples, following approval by the Cantonal Ethical Committee of Cantone Ticino.

**Isolation of monoclonal antibodies.** IgG<sup>+</sup> memory B cells were isolated from cryopreserved PBMC using CD22 microbeads (Miltenyi Biotec), followed by depletion of cells carrying IgM, IgD and IgA by cell sorting, and immortalized with EBV and CpG in multiple replicate wells as previously described<sup>14</sup>. Culture supernatants were tested for their ability to neutralize HRSV A2 infection in a micro-neutralization assay. Positive cultures were collected and expanded. From positive cultures the V<sub>H</sub> and V<sub>L</sub> sequences were retrieved by RT-PCR and cloned into human IgG1 and Igκ or Igλ expression vectors (provided by M. Nussenzweig) essentially as described<sup>31</sup>. Monoclonal antibodies were produced by transient transfection of 293 Freestyle cells (Invitrogen). Supernatants from transfected cells were collected after 7 days and IgG were affinity purified by protein A chromatography (GE Healthcare) and desalted against PBS.

**Inhibition of binding assay on HRSV infected cells.** Palivizumab and 101F were biotinylated using the EZ-Link NHS-PEO solid phase biotinylation kit (Pierce). Labelled antibodies were tested for binding to HRSV infected Hep-2 cells to determine the optimal concentration of each antibody to achieve 70–80% maximal binding. The biotin-labelled antibodies were then used as probes to assess, by flow cytometry, whether their binding (measured using fluorophore-conjugated streptavidin) was inhibited by pre-incubation of HRSV-A2-infected cells with a 50-fold excess of homologous or heterologous unlabelled antibodies.

**Virus neutralization.** Culture supernatants were analysed for the presence of neutralizing antibodies using a microneutralization assay based on infection of Hep-2 cells by HRSV strain A2 or of LLC-MK2 cells by HMPV strain A1/6621/01. Neat supernatants were incubated with 50–100 TCID<sub>50</sub> of viruses for 1 h at room temperature before addition of Hep-2 or LLC-MK2 target cells which were incubated for 6 or 8 days, respectively. Viable cells were detected using the WST-1 reagent (Roche). A secondary screening was performed to test whether the isolated monoclonal antibodies were able to neutralize both HMPV and HRSV. The IC<sub>50</sub>

was determined using the microneutralization assay described above with 100 TCID<sub>50</sub> of virus and viral infection was measured on day 3 or 4 (depending on the strain) by indirect immunofluorescence using an automated Pathway 855 analyser (BD). IC<sub>50</sub> values were calculated by interpolation of neutralization curves fitted with a 4-parameter nonlinear regression with a variable slope. Neutralization of PVM was measured by incubating serial dilutions of the antibodies with 1,000 TCID<sub>50</sub> that were then added on BHK cells. Three days after the cells were detached, fixed, stained with labelled anti-PVM antibodies and analysed by flow cytometry. IC<sub>50</sub> titres were measured as above.

**Antibody variants.** Germline sequences of the isolated antibodies were determined with reference to the IMGT database<sup>32</sup>. Antibody variants in which single or multiple mutations were reverted to the germline sequence were produced either by gene synthesis (GenScript) or by site-directed mutagenesis (Promega) and confirmed by sequencing. Palivizumab, 101F, 234 and D25 V<sub>H</sub> and V<sub>L</sub> sequences were synthesized by GenScript and the corresponding antibodies were expressed by transient transfection. Expression vectors for human IgG1 were engineered to introduce the L234A, L235A mutations (LALA variant lacking FcR and C1q binding).

**Transient expression of recombinant HRSV and HMPV F protein.** The full length F genes of HRSV A2 strain and HMPV NL/1/99 (AAS92882) were codon-optimized for eukaryotic cell expression and cloned into the pCMV1 vector (Genlantis). F protein mutants were generated by site-directed mutagenesis (Promega) and tested in transiently transfected 293T cells. Two days after transfection, cells were collected, fixed and permeabilized with saponin for immunostaining with MPE8 or palivizumab. Binding of antibodies to transfected cells was analysed using a Becton Dickinson FACSCanto2 (BD Biosciences) with FlowJo software (TreeStar).

**Construction of the F protein amino acid sequence dendrogram.** Sequences were aligned using the ClustalW program integrated in the MEGA version 5.0 package<sup>33</sup>. The best-fitting nucleotide substitution model was estimated by means of a hierarchical likelihood ratio test implemented in MEGA. The phylogenetic tree was inferred using the neighbour-joining distance method, with genetic distances calculated under the Tamura-Nei model with gamma-distributed rates among sites (T93+Γ). The reliability of the neighbour-joining tree was estimated using 1,000 bootstrap replications.

**Western blotting.** Hep-2 cells were infected with HRSV and LLC-MK2 cells with MPV, lysed with CHAPS and separated on SDS-PAGE gel under reducing or non-reducing conditions. Proteins were then transferred on a PVDF (polyvinylidene difluoride) membrane, which was then incubated with either MPE8, palivizumab or 234 antibodies. Antibody binding was detected with an anti-human HRP-conjugated antibody in combination with the enhanced chemiluminescence (ECL) western blotting detection reagent (GE Healthcare) and LAS4000 charge-coupled device (CCD) camera system.

**Evaluation of the prophylactic and therapeutic efficacy in mice infected with HRSV, HMPV or PVM.** Groups of 6–8-week-old female BALB/c mice or 129S6/svEv-Stat1-deficient mice (Taconic) were injected intravenously with purified antibodies at concentrations varying from 0.12 to 30 mg per kg. In both prophylactic and therapeutic experiments with HRSV or HMPV, mice were deeply anaesthetized with isoflurane and four hours later challenged intranasally (i.n.) with HRSV A/A2/61 or HMPV A/6621/01 containing  $0.7 \times 10^6$ – $1 \times 10^6$  p.f.u. In the therapeutic setting, mice received the antibodies 1, 2, or 3 days after infection. On day 5, mice were killed, lungs were isolated and homogenized for viral RNA extraction. RT-PCR on viral RNA was performed using a SYBR Green RT-PCR kit (QuantiFast). qPCR was performed on samples with β-actin and HRSV or HMPV specific primers using a CFX96 Real Time PCR Detection System (Bio-Rad). Lung homogenates were also analysed for the presence of mouse inflammatory cytokines TNF, IL6 and IFN-γ using a multiplex bead array system (CBA Flex Sets, BD). In the lethal model of PVM infection BALB/c mice were infected (i.n.) as above with  $3 \times \text{MLD}_{50}$  of PVM strain 15 (2,000 p.f.u.) and treated prophylactically or therapeutically at the indicated time points. Mice were monitored daily for survival and weight loss until day 14 post-infection. Animals that lost more than 25% of their initial body weight were euthanized in accordance with our animal study protocol. In addition, morbidity score (1, healthy; 2, ruffled fur at neck; 3, piloerection and difficulty breathing, less alert; 4, lethargic with laboured breathing; 5, premonitory, with emaciation and cyanosis; or 6, death) was obtained daily. To measure viral replication in lungs, PVM-challenged mice were killed on days 3, 4, 5 or 7 days to collect lung tissues. PVM lung titres were determined by qPCR using PVM specific primers as above. Animal experiments were approved by the Cantonal Authorities of Cantone Ticino.

**Selection of escape mutants.** A fixed amount of HRSV A/Long/56 or HMPV A/6621/01 virus containing  $16 \times 10^6$  p.f.u. was incubated with a constant amount ( $0.5 \text{ mg ml}^{-1}$ ) of antibody MPE8 or palivizumab for 1 h at 37 °C. The incubated mixture was then adsorbed by Hep-2 or LLC-MK2 cells for 4 h at  $10,000 \text{ p.f.u.}$  per well in 96-multiwell plates. Infected cells were then washed twice with PBS and

replenished with fresh medium containing  $50 \mu\text{g ml}^{-1}$  of selecting antibodies. The cytopathic effect (CPE) on infected cells was monitored for up to 6 days after infection. Supernatants from CPE-positive wells were collected for a subsequent round of infection. The sequences of the F proteins from monoclonal antibody resistant mutants (MARMS) were determined as above.

**Epitope mapping using structured peptides.** Overlapping linear 15-mer and cyclized 15-mer peptides based on the F protein of HRSV strain A2 (residues 26–516) were synthesized on polypropylene support (minicards) and were tested for reactivity with MPE8 as described<sup>34,35</sup>. Peptide binding analysis was carried out by an enzyme-linked immunosorbent assay (ELISA)-based format in which the coloured substrate was quantified with a CCD camera and an image processing system. The values mostly ranged from 0 to 3,000, a log scale similar to 1 to 3 of a standard 96-well plate ELISA-reader.

**Cloning, expression and purification of recombinant HRSV F protein.** A DNA construct encoding the F ectodomain (GenBank accession code FJ614814) (HRSV F residues 1 to 513, with fusion peptide residues 137 to 146 deleted) with a C-terminal 'foldon', tobacco etch virus (TEV) site, His tag and StreptagII was codon-optimized for mammalian expression (GenScript) and cloned into the pcDNA3.1 vector (Invitrogen). The stabilized pre-fusion form of HRSV F protein was designed by substitution of the residues Leu 481, Asp 489, Ser 509 and Asp 510 by cysteines as described<sup>18</sup> in the presence of the stabilizing bacteriophage T4 fibrin 'foldon' sequence. Both proteins were expressed by transient transfection of 293F cells. Cell supernatants were loaded on affinity column  $\text{Ni}^{2+}$ -nitrilotriacetic acid (NTA) resin (GE Healthcare). After NTA purification, all samples were run on StrepTactin II chromatography (Qiagen). Finally, proteins were further purified by size-exclusion chromatography (Superdex 200 from GE Healthcare) with PBS as a running buffer.

**Surface plasmon resonance.** SPR experiments were performed on a ProteOn-XPR36 instrument (Bio-Rad). Antibodies were immobilized on a ProteOn GLC sensor chip surface through amine coupling at 2,000 response units (RU) and a blank surface with no antibody was created under identical coupling conditions for use as a reference. HRSV F protein was injected at a flow rate of  $100 \mu\text{l min}^{-1}$ , at concentrations of 200, 100, 75, 50 and 25 nM in different sensor channels. The data were processed using Proteon software and double referenced by subtraction of the blank surface and buffer-only injection before local fitting of the data. Binding curves were locally fitted to the two states model, to perform a  $K_d$  calculation. For competition experiment, antibodies were covalently coupled to a GLC sensor chip

at 2,000 RU. HRSV F protein was injected at 200 nM and the competing antibody was injected after 20 s of dissociation time.

**HPLC and SDS-PAGE analysis.** The proteins were applied onto a TSK G3000 column equilibrated with PBS. The elution volume was determined by monitoring the ultraviolet absorption at 280 nm. The apparent molecular weight of the different proteins was determined from a calibration curve with protein standards (GE Healthcare). All HRSV F proteins eluted as a single peak with an apparent mass of 180 kilodaltons, consistent with a trimeric structure. SDS-PAGE was performed on NuPAGE Novex 4–12% Bis-Tris gel (Invitrogen) and the proteins were visualized using Imperial Blue (Thermo Scientific).

**CD spectropolarimetry.** Circular dichroism (CD) spectra from F proteins ( $0.5 \text{ mg ml}^{-1}$  in 10 mM  $\text{NaPO}_4$ , pH 7.4) were recorded on a Chirascan spectropolarimeter (Applied Photophysics) over the wavelength range of 195 to 260 nm at a band width of 1 nm, step size of 0.5 nm and 1 s per step. The spectra in the far-ultraviolet region required an average of five scans and were subtracted from blank spectra performed with buffer. Thermal denaturation was monitored by performing scans at intervals of  $5^\circ\text{C}$ , after equilibration for 2 min at each temperature. Data were fitted to a simple first order curve. The values of  $\Delta\text{A}222$  are represented on the y axis as the percentage of the values recorded at  $20^\circ\text{C}$ .

**Electron microscopy.** HRSV F protein sample ( $100 \mu\text{g ml}^{-1}$  in PBS) was absorbed onto a 400-mesh carbon-coated grid (and stained with 0.75% uranyl formate). A Philips CM 12 microscope, operated at 80 kV, was used to analyse the samples. Micrographs were taken at  $\times 65,000$  magnification.

**Statistics.** Student's two-tailed paired *t*-tests were used for statistical comparisons; *P* values of 0.05 or less were considered significant.

31. Tiller, T. *et al.* Efficient generation of monoclonal antibodies from single human B cells by single cell RT-PCR and expression vector cloning. *J. Immunol. Methods* **329**, 112–124 (2008).
32. Lefranc, M.-P. *et al.* IMGT, the international ImMunoGeneTics information system. *Nucleic Acids Res.* **37**, D1006–D1012 (2009).
33. Tamura, K. *et al.* MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28**, 2731–2739 (2011).
34. Timmerman, P., Puijk, W. C. & Melen, R. H. Functional reconstruction and synthetic mimicry of a conformational epitope using CLIPS technology. *J. Mol. Recognit.* **20**, 283–299 (2007).
35. Sliotstra, J. W., Puijk, W. C., Ligtoet, G. J., Langeveld, J. P. & Melen, R. H. Structural aspects of antibody-antigen interaction revealed through small random peptide libraries. *Mol. Divers.* **1**, 87–96 (1996).



# The structural mechanism of KCNH-channel regulation by the eag domain

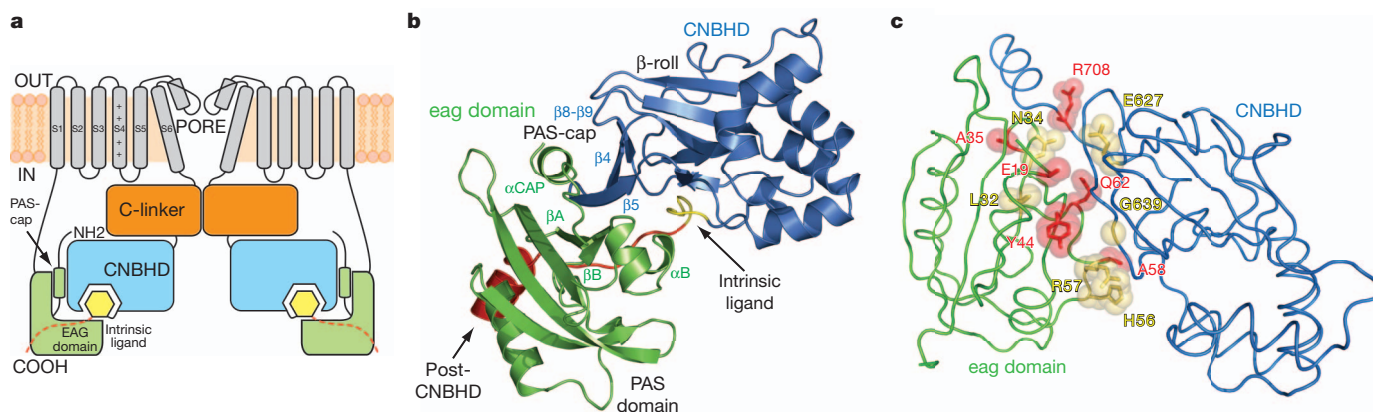
Yoni Haitin<sup>1</sup>, Anne E. Carlson<sup>1</sup> & William N. Zagotta<sup>1</sup>

The KCNH voltage-dependent potassium channels (*ether-à-go-go*, EAG; EAG-related gene, ERG; EAG-like channels, ELK) are important regulators of cellular excitability<sup>1–3</sup> and have key roles in diseases such as cardiac long QT syndrome type 2 (LQT2)<sup>4</sup>, epilepsy<sup>5</sup>, schizophrenia<sup>6</sup> and cancer<sup>7</sup>. The intracellular domains of KCNH channels are structurally distinct from other voltage-gated channels. The amino-terminal region contains an eag domain, which is composed of a Per-Arnt-Sim (PAS) domain and a PAS-cap domain<sup>8</sup>, whereas the carboxy-terminal region contains a cyclic nucleotide-binding homology domain (CNBHD), which is connected to the pore through a C-linker domain. Many disease-causing mutations localize to these specialized intracellular domains, which underlie the unique gating and regulation of KCNH channels<sup>9</sup>. It has been suggested that the eag domain may regulate the channel by interacting with either the S4–S5 linker or the CNBHD<sup>8,10</sup>. Here we present a 2 Å resolution crystal structure of the eag domain–CNBHD complex of the mouse EAG1 (also known as KCNH1) channel. It displays extensive interactions between the eag domain and the CNBHD, indicating that the regulatory mechanism of the eag domain primarily involves the CNBHD. Notably, the structure reveals that a number of LQT2 mutations at homologous positions in human ERG, in addition to cancer-associated mutations in EAG channels, localize to the eag domain–CNBHD interface. Furthermore, mutations at the interface produced marked effects on channel gating, demonstrating the important physiological role of the eag domain–CNBHD interaction. Our structure of the eag domain–CNBHD complex of mouse EAG1 provides unique insights into the physiological and pathophysiological mechanisms of KCNH channels.

KCNH channels have two large intracellular regions that underlie the specialized gating and regulation of this channel family<sup>9</sup> (Fig. 1a). The CNBHD in the C-terminal region shares similarity with the cyclic

nucleotide-binding domains of cyclic nucleotide-gated (CNG) and hyperpolarization-activated cyclic nucleotide-modulated (HCN) channels<sup>11,12</sup>. The CNBHD of KCNH channels, however, does not bind cyclic nucleotides and regulates KCNH channel gating in a cyclic nucleotide-independent manner<sup>13</sup>. The N-terminal region contains an eag domain, not present in CNG and HCN channels, that also has a key role in channel gating. Composed of a 110-amino-acid PAS domain preceded by a highly conserved 25-amino-acid PAS-cap (Fig. 1a and Supplementary Fig. 1), the eag domain regulates activation and inactivation in EAG1, and deactivation in human (h)ERG1 (also known as KCNH2) channels<sup>9,14</sup>. Moreover, many of the human mutations that cause LQT2 and cancer localize to the conserved eag domain<sup>15–17</sup>. In addition, high expression of a primate-specific isoform of ERG (KCNH2-3.1), missing most of the eag domain, was shown to correlate with risk for schizophrenia<sup>6</sup>. The mechanism underlying KCNH channel regulation by the eag domain remains unclear. Two competing hypotheses have emerged. Some evidence suggests that the eag domain interacts with the S4–S5 linker and directly regulates the movement of the S4 voltage sensor<sup>10,18</sup>. By contrast, other evidence favours the CNBHD as the interaction partner of the eag domain<sup>9,14,19</sup>. Using fluorescence anisotropy we found that the eag domain interacts directly with the CNBHD with an affinity of  $13.2 \pm 2.3 \mu\text{M}$  (Supplementary Fig. 2).

To uncover the mechanism for how the eag domain regulates the channel, we solved the X-ray crystal structure of the eag domain–CNBHD complex to 2.0 Å resolution (Supplementary Table 1). There were eight chains in the asymmetric unit: four eag domains and four CNBHDs. The spatial arrangement of the eag domains and CNBHDs within the asymmetric unit showed three different eag domain–CNBHD interactions, here termed interactions A, B and C (Supplementary Fig. 3). Interaction A, however, had considerably greater buried surface area than the other interactions (Supplementary Table 2) and was confirmed



**Figure 1 | Structure of the eag domain–CNBHD complex of mEAG1.**

**a**, Cartoon of a cross section of a KCNH channel. Transmembrane domains are in grey, the N-terminal eag domains are in green, the C-linkers are in orange and the CNBHD domains are in blue. The intrinsic ligand motifs are highlighted in yellow, and the post-CNBHD in red. **b**, Structure of the eag

domain–CNBHD complex. Colour corresponds to panel **a**. **c**, Disease-causing mutations at the interface of the eag domain–CNBHD complex. LQT2 mutations are shown in yellow. Cancer-associated mutations are shown in red. Y44 is involved in both LQT2 and cancer and is shown in red.

<sup>1</sup>Department of Physiology and Biophysics, University of Washington School of Medicine, Seattle, Washington 98195, USA.

by mutagenesis experiments in the intact channel (see below). Therefore interaction A was designated as the biological unit.

Figure 1b shows the structure of the eag domain–CNBHD complex in the biological unit. The structure of the CNBHD (residues 552–720) includes the last 35 amino acids of the C-linker, the CNBHD, and 22 amino acids of the post-CNBHD region (Fig. 1b and Supplementary Fig. 1). The structure of the eag domain (residues 16–136) includes the PAS domain and the last 10 amino acids of the PAS-cap domain. These structures share many similarities to structures of CNBHD and eag domains crystallized alone. The CNBHD is similar to the isolated structures of mouse (m)EAG1 (root mean squared deviation (r.m.s.d.) = 0.53 Å)<sup>12</sup>, mosquito ERG (r.m.s.d. = 1.05 Å)<sup>20</sup> and zebrafish ELK (r.m.s.d. = 1.39 Å)<sup>11</sup>. The PAS domain is similar to the isolated structures of mEAG1 (r.m.s.d. = 0.32 Å)<sup>21</sup>, hERG1 (r.m.s.d. = 0.71 Å)<sup>8,21</sup> and fruitfly ELK (r.m.s.d. = 0.62 Å)<sup>21</sup>, and also to the solution nuclear magnetic resonance (NMR) structures of hERG1 (refs 19, 22).

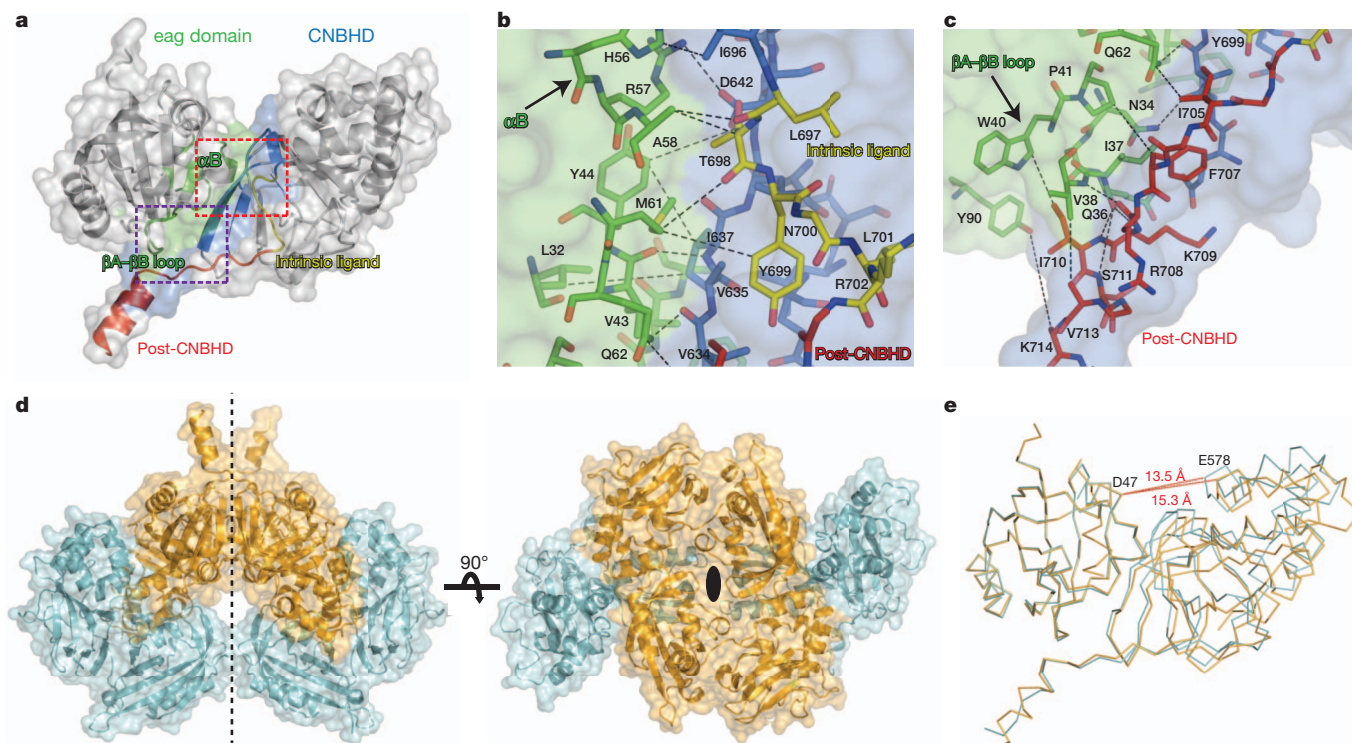
The interface between the eag domain and CNBHD in the co-crystal structure, however, is entirely novel. When we mapped the positions of disease-causing mutations in KCNH channels onto the structure of the eag domain–CNBHD complex, an unexpected trend emerged: many LQT2 mutations (at homologous positions in hERG1) and cancer-associated mutations (in hEAG1) are located at the interface between the eag domain and CNBHD (Fig. 1c and Supplementary Fig. 4). These mutations, which alter channel gating as well as trafficking to the membrane<sup>9,14</sup>, result in prolongation of the cardiac action potential (in hERG1) and were documented in several primary cancer types (in hEAG1)<sup>15</sup>. For example, position N34 in the eag domain, which forms hydrogen bonds with V634 of the CNBHD, was shown to cause LQT2 (hERG1(N33T)). In addition, Y44 of the eag domain forms an interaction with I637 and G639 of the CNBHD. Y44H in hEAG1 is a mutation known to be correlated with large intestine carcinoma<sup>15</sup>, and a mutation at the equivalent position in hERG1 (Y43C) causes LQT2 (ref. 23)

(Fig. 1c). G639 is homologous to hERG1(G800W), a position known to cause LQT2 (ref. 23). These observations suggest that perturbation of the interaction between the eag domain and the CNBHD underlies part of the pathophysiology of KCNH-related diseases. Moreover, the clustering of so many mutations at the eag domain–CNBHD interface is consistent with this interface found in the crystal representing the physiological architecture of KCNH channels.

The interface between the eag domain and the CNBHD is extensive, with an average buried solvent-accessible surface area of ~1,400 Å<sup>2</sup> (Supplementary Table 2). The interface consists of three sub-regions: (1) the intrinsic ligand motif of the CNBHD interacts with the  $\alpha$ B-helix in the PAS domain; (2) the  $\beta$ A and  $\beta$ B strands of the PAS domain interact with the post-CNBHD segment of the CNBHD; and (3) an amphipathic helix ( $\alpha$ CAP) in the PAS-cap domain forms an interaction with the  $\beta$ -roll of the CNBHD (Fig. 1a, b).

The direct interaction of the PAS domain (residues 57–61) with the intrinsic ligand motif of the CNBHD was particularly surprising (Fig. 2a, b). Previous structures of isolated CNBHD from mEAG1 (ref. 12) and zebrafish ELK<sup>11</sup> showed that a short loop (residues 697–701) following the  $\alpha$ C helix occupies the space filled by cyclic nucleotides in canonical cyclic nucleotide-binding domains (Supplementary Fig. 1). This short loop was therefore called the ‘intrinsic ligand’. Mutations in this intrinsic ligand were shown to regulate channel activity, shifting the voltage dependence and kinetics of activation and deactivation. Interestingly, mutations in the eag domain of KCNH channels result in similar alterations<sup>9,14,16</sup>. The interaction between the eag domain and the intrinsic ligand, and the functional similarities between mutations in these domains, suggests that the intrinsic ligand motif may be a critical determinant of the regulation of KCNH channels by the eag domain.

Another surprising site of interaction of the eag domain is the post-CNBHD region (beginning at residue 702). This region, immediately C-terminal to the intrinsic ligand, has been shown to mediate the regulation



**Figure 2 | The various interfaces between the eag domain and CNBHD, and different complex types in the asymmetric unit.** a–c, Cartoon and surface representation of the eag domain–CNBHD complex (a). Interaction surface of the eag domain and the CNBHD is shown in green and blue, respectively. The intrinsic ligand and post-CNBHD region are in yellow and red, respectively. The non-interacting parts of the structure are in grey. Red and purple rectangles

define regions detailed in b (intrinsic ligand) and c (post-CNBHD region), respectively. d, Cartoon and surface representation of the asymmetric unit complexes related by non-crystallographic symmetry in the same conformation. Two-fold symmetry axis indicated. e, Comparison of the structures of the two complex types, colour coded as in d. Dashed red lines show distances between D47 of the eag domain and E578 of the CNBHD.



of KCNH channel function by a variety of cellular signalling events, including phosphorylation, and interaction with kinases, integrins and  $\text{Ca}^{2+}$ -calmodulin<sup>24–26</sup>. The eag domain–CNBHD structure shows that the post–CNBHD region interacts with the  $\beta\text{A}$ – $\beta\text{B}$  loop of the eag domain (Fig. 2a, c and Supplementary Fig. 1). The interaction between the eag domain and the post–CNBHD region suggests that KCNH channel regulation through the post–CNBHD region may involve its interaction interface with the eag domain.

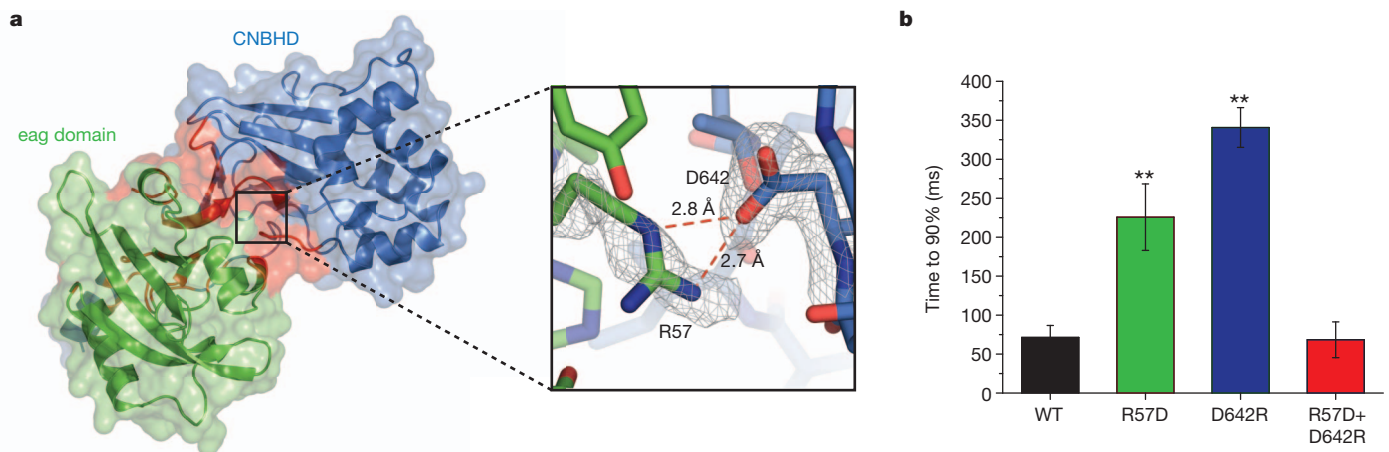
The crystal contains two different conformations of the eag domain–CNBHD complex. The asymmetric unit is composed of four copies of the complex, related by a two-fold non-crystallographic symmetry (Fig. 2d). This results in two pairs of the complex ('type I' and 'type II' complexes) that are in a similar conformation within each pair (average r.m.s.d. = 0.16 Å), but a different conformation and B-factor distribution between pairs (r.m.s.d. = 0.96 Å) (Fig. 2d and Supplementary Fig. 5). Alignment of the two conformations reveals that the angle between the eag domain and the CNBHD differs by 5° and that the distance between the two domains differs by 1.8 Å (between D47 of the eag domain and E578 of the CNBHD) (Fig. 2e and Supplementary Video 1). In addition, the  $\alpha\text{P}$  helix (residues 649–651; Supplementary Fig. 1) undergoes a substantial rearrangement between the two conformations, possibly due, at least in part, to variation in the local crystal contacts (Fig. 2e and Supplementary Video 1). These two conformations, if present in the intact channel, suggest that the eag domain–CNBHD complex is not static and represent a possible rearrangement associated with the gating regulation of the channel by the eag domain.

To determine the functional role of the eag domain–CNBHD interaction, we mutated the interaction interface between the eag domain and the CNBHD in intact mEAG1 channels. The interface contains a salt bridge between R57 in the  $\alpha\text{B}$  helix of the PAS domain and D642 in the  $\beta\text{6}$  strand of the CNBHD (Fig. 3a and Supplementary Fig. 4). This salt bridge is present in both conformations of the biological unit and is not present in the other eag domain–CNBHD interactions in the crystal. Notably, R57 and D642 are conserved throughout the KCNH family (Supplementary Fig. 1), and mutation of the R57-equivalent position in hERG1 (hERG1(R56Q)) causes LQT2 (ref. 16). mEAG1 channels with mutations in the salt bridge were expressed in *Xenopus* oocytes and their currents were recorded. Individually the eag domain R57D mutation and the CNBHD D642R mutation each produced a slowing of activation with voltage steps to +50 mV (Supplementary Fig. 6). This resulted primarily from a new slow component to the activation time course and produced a three- to fivefold delay of the time to 90% current amplitude when compared to wild-type channels (Fig. 3b) (R57D,  $225 \pm 42$  ms; D642R,  $340 \pm 25$  ms; wild type,  $71 \pm 15$  ms;

$P < 0.01$ ;  $n = 9–10$ ). No changes in other biophysical properties of the channels were noted (Supplementary Table 3). Importantly, in the double charge reversal mutant, R57D + D642R, the time course of channel activation was similar to wild type ( $68 \pm 23$  ms,  $P > 0.05$ ,  $n = 6$ ), suggesting that the salt bridge was restored (Fig. 3b). These results indicate that the R57–D642 salt bridge is present in the intact channel, consistent with interaction A being the biological unit. In addition, these results indicate that the eag domain regulates the activation of mEAG1 channels by interaction with the CNBHD.

A third important region of interaction with the CNBHD is with the PAS-cap domain. The PAS-cap domain constitutes the first 25 amino acids of the eag domain. This domain is highly conserved within the KCNH family (Supplementary Fig. 1) and has been shown to be crucial for proper activation and inactivation gating in KCNH channels<sup>9,27</sup>. Notably, our structure of the eag domain–CNBHD complex reveals that an amphipathic helix ( $\alpha\text{CAP}$ ) of the PAS-cap domain of mEAG1 interacts directly with the CNBHD (Fig. 4a). The structure shows that  $\alpha\text{CAP}$  is nestled between a hydrophobic patch of the PAS domain and the  $\beta\text{4}$ – $\beta\text{5}$  strands and  $\beta\text{8}$ – $\beta\text{9}$  loop of the CNBHD. The structure of the PAS-cap domain was previously assessed using NMR of isolated eag domain from hERG1 (refs 10, 19, 22). Structural alignment reveals that the PAS-cap domain in our eag domain–CNBHD complex is in a very different orientation from the NMR structures (Fig. 4b). The orientation, however, is very similar to recent X-ray crystallography structures of isolated eag domains from hERG1 and ELK<sup>21</sup>. We propose that the spatial organization of the PAS-cap domain within the eag domain–CNBHD complex recapitulates the native conformation of this domain in the intact channel, and that the PAS-cap domain exerts its functional effects through its interaction with the CNBHD.

Although the PAS-cap domain itself does not contain any known LQT2 mutations and only one cancer-associated mutation (hEAG1(E19D)), its surroundings are rich in pathological loci, including hERG1 LQT2 mutation E788D<sup>28</sup> (E627 in mEAG1) in the  $\beta\text{4}$  strand of the CNBHD<sup>9,15</sup>. These mutations may alter the gating properties of the channel by destabilizing the PAS-cap domain interaction with the CNBHD. To test this hypothesis, we measured the currents from mEAG1 channels mutated in the PAS-cap domain (R7A–R8A and R7E–R8E) and in the CNBHD region where the PAS-cap is bound (E627A and E627R). These positions are highly conserved within the KCNH family (Supplementary Fig. 1). The voltage dependence of activation of the PAS-cap mutations R7A–R8A and R7E–R8E were significantly shifted to more depolarized potentials, compared to wild-type channels (Fig. 4c and Supplementary Table 4). The CNBHD mutations also showed a robust depolarizing shift, similar to that of the R7E–R8E mutant (Fig. 4c

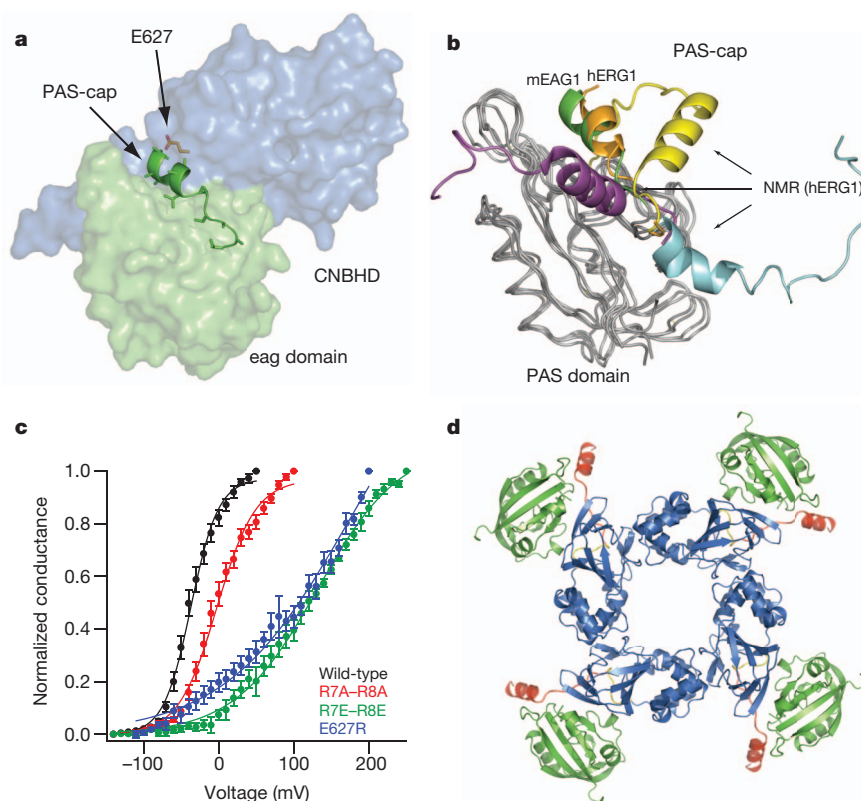


**Figure 3 | A salt bridge between the eag domain and the CNBHD.**

**a**, Cartoon and surface representation of the primary interaction interface (interaction A) of the eag domain–CNBHD complex (left). Eag domain is shown in green, CNBHD in blue and interface residues in red. A detailed

illustration of the R57–D642 salt bridge is shown on the right, overlaid with  $2F_o - F_c$  composite 'omit' map contoured at  $1\sigma$  (grey mesh). **b**, Summary of time to 90% maximum current amplitude for wild-type (WT) and mutants (error bars indicate  $\pm$  s.e.m.;  $n = 6–10$  cells;  $P < 0.005$ , Student's *t*-test).





**Figure 4 | PAS-cap interaction with the CNBHD, and predicted tetramer of the eag domain–CNBHD complex.** **a**, Cartoon representation of the PAS-cap region. The PAS domain and the CNBHD are shown as surface (green and blue, respectively) with the PAS-cap regions shown as cartoon and position E627 as orange sticks. **b**, Alignment of the eag domain (green) with representative hERG1 eag domain structures. The X-ray structures of the eag domain–CNBHD complex of mEAG1 and the isolated eag-domain of hERG1 (PDB

code 4HP9) are shown as green and orange cartoons, respectively, and the NMR structures are in yellow (PDB code 2L0W), cyan (PDB code 2L4R) and purple (PDB code 2L1M). The PAS domains are represented as grey ribbons. **c**, Normalized conductance plotted as a function of the test voltages for mEAG1 wild-type and mutant channels (error bars indicate  $\pm$  s.e.m.;  $n = 3$ –68 cells;  $P < 0.001$ , Student's *t*-test). **d**, Alignment of the eag domain–CNBHD complex with the CNBD of mHCN2 (PDB 1Q50), colour coded as in Fig. 1b.

and Supplementary Table 4). Together, these results suggest that the interaction between the PAS-cap domain and the CNBHD regulates KCNH channel gating.

To model the orientation of the eag domain–CNBHD complex in the context of a four-fold symmetrical ion channel, we aligned the CNBHD of the complex with the homologous C-linker/CNBD of the mHCN2 channel (PDB code 1Q50), which shares 35% sequence similarity with the CNBHD of mEAG1 (Supplementary Fig. 1). The C-linker/CNBD in the mHCN2 structure assembles as a tetramer with a four-fold axis of rotational symmetry, in which the C-linker domains form most of the inter-subunit interactions<sup>29</sup>. Although our structure of the mEAG1 CNBHD lacks most of C-linker, the alignment resulted in a concentric arrangement of the eag domain–CNBHD complex, where the eag domains are positioned at the periphery of the C-linker/CNBD tetramer (Fig. 4d). The post-CNBHD region points away from the centre and may be accessible for interaction with other proteins, such as kinases and  $\text{Ca}^{2+}$ -calmodulin, thought to regulate KCNH channels through sites in the post-CNBHD<sup>26</sup>. Mutagenesis studies in hERG1 channels suggest that this arrangement may exhibit domain swapping in which the eag domain of one subunit interacts with the CNBHD of a neighbouring subunit<sup>9</sup>.

The eag domain and the CNBHD have been shown to be important regulatory domains of the KCNH channel family. Mutations found in these regions lead to LQT2 and are associated with cancer. The structure of the eag domain–CNBHD complex of mEAG1 demonstrates that the interaction interface between the two domains serves as a hot-spot for many of these mutations and provides a framework to better understand the mechanisms of gating and regulation of KCNH channels. Furthermore, understanding the molecular interactions within the

KCNH channel complex, and their disruption in disease, may enable design of novel therapies for these devastating conditions.

## METHODS SUMMARY

The mouse EAG1 eag domain (residues 6–136) and CNBHD (residues 552–724) were subcloned into a pETM11 vector containing an N-terminal hexa-histidine affinity tag followed by a tobacco etch virus (TEV) cleavage site. The proteins were expressed in BL21(DE3) *Escherichia coli* cells as described previously<sup>11</sup>, purified on a  $\text{Ni}^{2+}$  affinity column and then on an ion-exchange (CNBHD) or size-exclusion (eag domain) column, following an overnight cleavage with TEV at 4 °C. The proteins were concentrated to about 475  $\mu\text{M}$ , flash-frozen and stored at  $-80$  °C until use. The eag domain and the CNBHD were mixed in a 1:1 molar ratio before experimentation. Crystals were grown at 20 °C using sitting-drop vapour diffusion by mixing a 2:1 (v/v) ratio of protein mixture and a reservoir solution containing 18% (w/v) PEG 3350, 1.8% (v/v) tacsimate, 10 mM  $\text{MnCl}_2$  and 90 mM HEPES, pH 7.5. The crystals diffracted to 1.995 Å and belong to space group  $P6_5$  (Supplementary Table 1 and Supplementary Fig. 7). The structure was solved by molecular replacement using the mEAG1 CNBHD (PDB accession 4F8A) as the search model. The final model was refined to a resolution of 1.995 Å with  $R_{\text{work}} = 0.167$  and  $R_{\text{free}} = 0.197$ . Electrophysiology recordings were performed using the inside-out configuration of the patch-clamp technique on *Xenopus* oocytes expressing wild-type and mutant channels as described previously<sup>11</sup>.

**Full Methods** and any associated references are available in the online version of the paper.

Received 30 April; accepted 18 July 2013.

Published online 25 August 2013.

- Warmke, J., Drysdale, R. & Ganetzky, B. A distinct potassium channel polypeptide encoded by the *Drosophila* eag locus. *Science* **252**, 1560–1562 (1991).
- Ganetzky, B., Robertson, G. A., Wilson, G. F., Trudeau, M. C. & Titus, S. A. The eag family of  $\text{K}^+$  channels in *Drosophila* and mammals. *Ann. NY Acad. Sci.* **868**, 356–369 (1999).

3. Bauer, C. K. & Schwarz, J. R. Physiology of EAG K<sup>+</sup> channels. *J. Membr. Biol.* **182**, 1–15 (2001).
4. Sanguinetti, M. C. & Tristani-Firouzi, M. hERG potassium channels and cardiac arrhythmia. *Nature* **440**, 463–469 (2006).
5. Zhang, X. *et al.* Deletion of the potassium channel Kv12.2 causes hippocampal hyperexcitability and epilepsy. *Nature Neurosci.* **13**, 1056–1058 (2010).
6. Huffaker, S. J. *et al.* A primate-specific, brain isoform of KCNH2 affects cortical physiology, cognition, neuronal repolarization and risk of schizophrenia. *Nature Med.* **15**, 509–518 (2009).
7. Camacho, J. *Ether à go-go* potassium channels and cancer. *Cancer Lett.* **233**, 1–9 (2006).
8. Morais Cabral, J. H. *et al.* Crystal structure and functional analysis of the hERG potassium channel N terminus: a eukaryotic PAS domain. *Cell* **95**, 649–655 (1998).
9. Zastina, A. S. & Trudeau, M. C. hERG potassium channel regulation by the N-terminal eag domain. *Cell. Signal.* **24**, 1592–1598 (2012).
10. Li, Q. *et al.* NMR solution structure of the N-terminal domain of hERG and its interaction with the S4–S5 linker. *Biochem. Biophys. Res. Commun.* **403**, 126–132 (2010).
11. Brelidze, T. I., Carlson, A. E., Sankaran, B. & Zagotta, W. N. Structure of the carboxy-terminal region of a KCNH channel. *Nature* **481**, 530–533 (2012).
12. Marques-Carvalho, M. J. *et al.* Structural, biochemical, and functional characterization of the cyclic nucleotide binding homology domain from the mouse EAG1 potassium channel. *J. Mol. Biol.* **423**, 34–46 (2012).
13. Brelidze, T. I., Carlson, A. E. & Zagotta, W. N. Absence of direct cyclic nucleotide modulation of mEAG1 and hERG1 channels revealed with fluorescence and electrophysiological methods. *J. Biol. Chem.* **284**, 27989–27997 (2009).
14. Stevens, L., Ju, M. & Wray, D. Roles of surface residues of intracellular domains of hERG potassium channels. *Eur. Biophys. J.* **38**, 523–532 (2009).
15. Bamford, S. *et al.* The COSMIC (Catalogue of Somatic Mutations in Cancer) database and website. *Br. J. Cancer* **91**, 355–358 (2004).
16. Chen, J., Zou, A., Splawski, I., Keating, M. T. & Sanguinetti, M. C. Long QT syndrome-associated mutations in the Per-Arnt-Sim (PAS) domain of hERG potassium channels accelerate channel deactivation. *J. Biol. Chem.* **274**, 10113–10118 (1999).
17. Splawski, I. *et al.* Spectrum of mutations in long-QT syndrome genes *KVLQT1*, *HERG*, *SCN5A*, *KCNE1*, and *KCNE2*. *Circulation* **102**, 1178–1185 (2000).
18. Wang, J., Trudeau, M. C., Zappia, A. M. & Robertson, G. A. Regulation of deactivation by an amino terminal domain in human *ether-à-go-go-related gene* potassium channels. *J. Gen. Physiol.* **112**, 637–647 (1998).
19. Muskett, F. W. *et al.* Mechanistic insight into human *ether-à-go-go-related gene* (hERG) K<sup>+</sup> channel deactivation gating from the solution structure of the EAG domain. *J. Biol. Chem.* **286**, 6184–6191 (2011).
20. Brelidze, T. I., Gianulis, E. C., Dimaio, F., Trudeau, M. C. & Zagotta, W. N. Structure of the C-terminal region of an ERG channel and functional implications. *Proc. Natl Acad. Sci. USA* **110**, 11648–11653 (2013).
21. Adaixo, R., Harley, C. A., Castro-Rodrigues, A. F. & Morais-Cabral, J. H. Structural properties of PAS domains from the KCNH potassium channels. *PLoS ONE* **8**, e59265 (2013).
22. Ng, C. A. *et al.* The N-terminal tail of hERG contains an amphipathic  $\alpha$ -helix that regulates channel deactivation. *PLoS ONE* **6**, e16191 (2011).
23. Napolitano, C. *et al.* Genetic testing in the long QT syndrome: development and validation of an efficient approach to genotyping in clinical practice. *J. Am. Med. Assoc.* **294**, 2975–2980 (2005).
24. Cherubini, A. *et al.* Human *ether-a-go-go-related gene* 1 channels are physically linked to  $\beta_1$  integrins and modulate adhesion-dependent signaling. *Mol. Biol. Cell* **16**, 2972–2983 (2005).
25. Sun, X. X., Hodge, J. J., Zhou, Y., Nguyen, M. & Griffith, L. C. The eag potassium channel binds and locally activates calcium/calmodulin-dependent protein kinase II. *J. Biol. Chem.* **279**, 10206–10214 (2004).
26. Ziechner, U. *et al.* Inhibition of human *ether à go-go* potassium channels by Ca<sup>2+</sup>/calmodulin binding to the cytosolic N- and C-termini. *FEBS J.* **273**, 1074–1086 (2006).
27. Wang, J., Myers, C. D. & Robertson, G. A. Dynamic control of deactivation gating by a soluble amino-terminal domain in hERG K<sup>+</sup> channels. *J. Gen. Physiol.* **115**, 749–758 (2000).
28. Tester, D. J., Will, M. L., Haglund, C. M. & Ackerman, M. J. Compendium of cardiac channel mutations in 541 consecutive unrelated patients referred for long QT syndrome genetic testing. *Heart Rhythm* **2**, 507–517 (2005).
29. Zagotta, W. N. *et al.* Structural basis for modulation and agonist specificity of HCN pacemaker channels. *Nature* **425**, 200–205 (2003).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank S. Cunnington and S. Camp for technical assistance. We thank the beamline staff at the Advanced Light Source (ALS). We also thank G. Flynn, J. Bankston, M. Puljung and S. Gordon for helpful discussions and critical reading of the manuscript. This work was supported by the National Institutes of Health (NIH) grant R01 EY010329 (W.N.Z.), The International Human Frontier Science Program Organization (HFSP) long-term postdoctoral fellowship LT-001025/2011 (Y.H.), and NIH grant F32 HL095241 (A.E.C.). The Berkeley Center for Structural Biology is supported in part by the National Institutes of Health, National Institute of General Medical Sciences, and the Howard Hughes Medical Institute. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the US Department of Energy under Contract No. DE-AC02-05CH11231.

**Author Contributions** Y.H. and W.N.Z. conceived the experiments. Y.H. performed the biochemical and crystallographic experiments and data analysis. A.E.C. conceived the PAS-cap-related experiments and performed all the electrophysiology experiments and data analysis. Y.H. and W.N.Z. wrote the manuscript.

**Author Information** Atomic coordinates and structure factors for the reported crystal structures have been deposited with the Protein Data Bank under accession code 4LLO. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to W.N.Z. (zagotta@uw.edu).

## METHODS

**Cloning, expression and purification.** Mouse EAG1 (Genbank accession codes NP034730.1 and GI:6754422) eag domain (residues 6–136) and CNBHD (residues 552–712, 552–724 and 552–752) were subcloned using 5' NcoI and 3' HindIII sites into a pETM11 vector containing an N-terminal hexa-histidine affinity tag followed by a tobacco etch virus protease (TEV) cleavage site. The internal NcoI restriction site in the CNBHD was removed via a silent mutation. Final CNBHD boundaries were selected based on small-scale expression and purification experiments (residues 552–724). All constructs had a GAM(G) cloning artefact sequence introduced at the N terminus. Full-length complementary DNA of mEAG1 was subcloned into the high-expression pGH19 vector for expression in *Xenopus* oocytes. The proteins were expressed in BL21(DE3) *Escherichia coli* cells as previously described<sup>11</sup>. Bacterial cultures were grown to mid-log phase and induced with 1 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) overnight at 18 °C. The cultures were spun down and re-suspended in 150 mM NaCl, 20 mM Tris-HCl and 1 mM tris(2-carboxyethyl)phosphine (TCEP), pH 8.0, containing 1 mM 4-(2-aminoethyl) benzenesulphonyl fluoride hydrochloride (AEBSF), 2.5  $\mu$ g ml<sup>-1</sup> DNaseI and cComplete protease inhibitor tablets (Roche). Cells were lysed with an EmulsiFlex C-5 homogenizer (Avestin) and the lysate was cleared by centrifugation at 131,000g for 45 min at 4 °C. The proteins were then purified on a Ni<sup>2+</sup> affinity resin column (HisTrap HP, GE Healthcare). The hexa-histidine tag was removed by TEV cleavage overnight at 4 °C. The CNBHD was further purified by anion exchange chromatography (HiTrap Q HP, GE Healthcare) after tenfold dilution with buffer containing 20 mM Tris-HCl, 1 mM TCEP, pH 8.0. The protein eluted with a shallow linear NaCl gradient as one well-separated peak. The TEV-digested eag domain was concentrated and further purified by gel filtration (Superdex 200, GE Healthcare) with buffer containing 150 mM NaCl, 20 mM Tris-HCl, 5 mM dithiothreitol (DTT), pH 8.0. Both proteins were concentrated to about 475  $\mu$ M using 10-kD molecular weight cut-off concentrators (Millipore), flash-frozen in liquid N<sub>2</sub> and stored at -80 °C until use. Fluorescein labelling of the endogenous cysteines of the eag domain (FL-eag) was done as described previously<sup>30</sup>. In brief, TEV-digested, gel-filtered, purified protein was incubated with 100  $\mu$ M fluorescein-5-maleimide (Invitrogen) for 1 h at 20 °C, in the above gel filtration buffer, without DTT. The labelled protein was separated from unincorporated dye by gel filtration and concentrated to 50  $\mu$ M.

**Fluorescence size-exclusion chromatography (FSEC)<sup>31</sup> and fluorescence anisotropy.** For FSEC experiments, FL-eag was mixed in a 1:1 molar ratio with purified CNBHD and the mixture was loaded on a Superdex 75 10/300 GL column (GE Healthcare) mounted on a high-performance liquid chromatography system (Shimadzu) with a fluorescence detector set for detection of fluorescein fluorescence. Fluorescence anisotropy was recorded using a Fluorolog 3 spectrophotometer (HORIBA, Jobin Yvon). In total, 100 nM FL-eag was placed in a quartz cuvette, and anisotropy was measured with increasing concentrations of CNBHD. Anisotropy experiments with 492 nm excitation and 514 nm emission were performed as described previously<sup>32</sup>. To estimate binding affinity, plots of the anisotropy versus total CNBHD concentration were fit using the following first-order reaction scheme:



$$\text{Anisotropy} = \alpha \left[ \frac{(R_t + K_d + L_t) - \sqrt{(-R_t - K_d - L_t)^2 - 4 \times R_t \times L_t}}{2} \right] + \beta$$

in which  $R$ ,  $L$  and  $RL$  are concentrations of the free receptor, free ligand, and receptor–ligand complex, respectively,  $R_t$  and  $L_t$  are total receptor and total ligand concentrations, respectively,  $K_d$  is the ligand-binding affinity, and  $\alpha$  and  $\beta$  are a scaling factor and an offset factor, respectively.

**Crystallization and structure determination.** Crystals of the eag domain–CNBHD complex were grown at 20 °C using sitting-drop vapour diffusion by mixing a 2:1 ratio (Mosquito, TTP LABTECH) of protein mixture and a reservoir solution, containing 18% (w/v) PEG 3350, 1.8% (v/v) Tacsimate, 10 mM MnCl<sub>2</sub> and 90 mM HEPES, pH 7.5. This condition produced crystals within 3 days, which grow to maximum size of about 300  $\mu$ m  $\times$  80  $\mu$ m  $\times$  80  $\mu$ m after 21 days. For diffraction data collection, crystals were immersed in liquid nitrogen after cryoprotection in 20% glycerol. Data were collected at 110 K on beamline 8.2.1 at the Advanced Light

Source (Lawrence Berkeley National Laboratory). Integration, scaling and merging of the diffraction data were done with the Mosflm program<sup>33</sup>. The crystals belonged to space group  $P6_5$  and the unit cell had dimensions of  $a = b = 162.38$  Å and  $c = 100.44$  Å with  $\alpha = \beta = 90^\circ$  and  $\gamma = 120^\circ$ . The structure was solved by molecular replacement using the programs Phaser<sup>34</sup> and Phenix<sup>35</sup> with the mEAG1 CNBHD (PDB accession 4F8A) as the search model. Data collection and refinement statistics are summarized in Supplementary Table 1. Each asymmetric unit contained eight protein subunits forming four eag domain–CNBHD complexes. Electron density was visible for most of the protein chains except for a few disordered terminal residues and the initial 10 and 23 residues of the eag domains of 'type I' and 'type II' dimers, respectively. Iterative model building and refinement were carried out in Phenix with manual adjustments using COOT<sup>36</sup>. The final model was refined to a resolution of 1.995 Å with  $R_{\text{work}} = 0.167$  and  $R_{\text{free}} = 0.197$ . The majority (98.2%) of the residues are in the most favoured region in the Ramachandran plot. All structural illustrations were prepared with PYMOL (<http://www.pymol.org>).

**Electrophysiology.** The cDNA encoding the mEAG1 channel in the pGH19 vector was kindly provided by G. Robertson (University of Wisconsin-Madison). The cRNA was transcribed using the T7 MEGascript kit (Ambion). *Xenopus laevis* oocytes were de-folliculated and injected with the cRNA as described previously<sup>37</sup>. After manual removal of the vitelline membrane, currents were recorded in the inside-out patch configuration<sup>38</sup> with an EPC-10 patch-clamp amplifier (HEKA Elektronik). Patch pipettes were pulled from borosilicate glass and had resistances of 0.4–0.8 M $\Omega$  after fire polishing. The intracellular (bath) and extracellular (pipette) solutions contained 130 mM KCl, 10 mM HEPES and 0.2 mM EDTA, pH 7.2. For the experiments examining the eag domain–CNBHD salt bridge (with the R57D and D642R mutations) the mEAG1 currents were elicited by applying a series of 1–2-s test pulses to voltages ranging from -120 to +50 mV in 10-mV increments from a holding potential of -100 mV, followed by a 0.5-s voltage pulse to -90 mV. For experiments examining an interaction between the PAS-cap and CNBHD (with mutations in residues R7–R8 and E627), the mEAG1 currents were elicited by a series of 50-ms test pulses to voltages ranging from -140 mV to +200 mV in 10-mV increments, from a holding potential of -100 mV, followed by a 200-ms voltage pulse to -100 mV. Currents were not leak subtracted. Data were acquired with Pulse software (HEKA Elektronik) and analysed with Igor (WaveMetrics).

To measure the conductance–voltage relationships for all electrophysiology experiments, peak tail current amplitudes at -90 or -100 mV, were normalized to the largest peak conductance amplitude. Plots of the normalized conductance versus the test voltage were fit with a Boltzmann function:

$$\frac{G}{G_{\text{max}}} = \frac{1}{1 + e^{-\left(\frac{V - V_{0.5}}{s}\right)}}$$

in which  $V$  represents the test voltage (mV),  $V_{0.5}$  is the midpoint activation voltage (mV) and  $s$  is the slope of the relation (mV).

30. Taraska, J. W., Puljung, M. C., Olivier, N. B., Flynn, G. E. & Zagotta, W. N. Mapping the structure and conformational movements of proteins with transition metal ion FRET. *Nature Methods* **6**, 532–537 (2009).
31. Kawate, T. & Gouaux, E. Fluorescence-detection size-exclusion chromatography for precrystallization screening of integral membrane proteins. *Structure* **14**, 673–681 (2006).
32. Rossi, A. M. & Taylor, C. W. Analysis of protein–ligand interactions by fluorescence polarization. *Nature Protocols* **6**, 365–387 (2011).
33. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
34. McCoy, A. J. Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr. D* **63**, 32–41 (2007).
35. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
36. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
37. Zagotta, W. N., Hoshi, T. & Aldrich, R. W. Gating of single Shaker potassium channels in *Drosophila* muscle and in *Xenopus* oocytes injected with Shaker mRNA. *Proc. Natl Acad. Sci. USA* **86**, 7243–7247 (1989).
38. Hamill, O. P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. J. Improved patch-clamp techniques for high-resolution current recording from cells and cell-free membrane patches. *Pflügers Arch.* **391**, 85–100 (1981).



# CAREERS

**FACULTY** Mentoring programme helps to improve grant and publication rates **p.451**

**SALARY** US holders of maths master's earn more than biologists or chemists **p.451**

**NATUREJOBS** For the latest career listings and advice [www.naturejobs.com](http://www.naturejobs.com)

DAVID SIMONDS



LAB LIFE

## Battle zone

*Disputes are bound to happen in high-pressure research environments. The key is knowing how to respond when they do.*

BY KAREN KAPLAN

**E**rica Sparkenbaugh was apprehensive when her principal investigator (PI) asked her to finish a paper that a more senior postdoc had started. But, as a recent arrival to the lab, she did not want to seem uncooperative. "I was kind of embarrassed, and I felt bad," remembers Sparkenbaugh, who researches inflammation and coagulation at the University of North Carolina at Chapel Hill. "I didn't want this to be the start of my new postdoc." Soon, she started hearing whispers that her colleague was upset about having his project snatched away. Worried about the future of their working relationship and about lab camaraderie, Sparkenbaugh approached her miffed lab-mate on a quiet evening at the

campus postdoctoral office. She explained that she had not intended to step on his toes. "I walked up to him and asked if he was OK with my finishing the paper," she says. "I wanted him to know that our PI had asked me to finish things up because he knew [the colleague] was busy with other projects."

The conversation proved fruitful. Her colleague's vexation evaporated, and he agreed that they should talk to their supervisor about a more collaborative approach. The three settled on a timeline for the postdocs to finish the paper together and submit it to a journal, and all went according to plan: the paper is now under review. "I think it was really helpful that I went to him directly," says Sparkenbaugh. "He and I have a really good working relationship now — it was worth those 10 minutes of sweating."

All sorts of discord can arise in the lab, tripping up the most well-meaning postdocs and graduate students. Differing expectations and ineffective or insufficient exchange of information are at the root of many clashes, say conflict-resolution experts. Younger researchers might be unclear about their own or their lab-mates' responsibilities, or might not fully understand how the lab functions. They might also communicate poorly, and unwittingly come across as truculent or confrontational.

Junior researchers can reduce the risk of serious, potentially career-marring conflicts through a few simple measures: gathering information before arriving at a lab; having a thorough chat with the principal investigator and other lab members; and remaining calm when disagreements arise. It can also help ►

► to consult a senior colleague, an administrator or a support organization such as the campus postdoc office — especially if a dispute seems too delicate or explosive to deal with alone. Ignoring a volatile situation, or letting it fester, will not make it disappear, and may put a researcher's good standing at risk.

## SETTING THE STAGE

As soon as graduate students or postdocs know that they will be working in a particular lab, they should aim to find out as much as possible about their supervisor and team. Useful information on a principal investigator's personality and management style can be unearthed by talking to his or her current or former lab members, suggests Charles Dwyer, an executive coach and an emeritus professor of education at the University of Pennsylvania in Philadelphia. "Find out what has happened with people who have worked with this professor before," he says. Such off-the-record, first-hand information is invaluable, agrees John Baldoni, an executive coach in Ann Arbor, Michigan. "You want to understand where your boss is coming from," he says.

Once armed with this intelligence, the junior researcher should arrange a chat with the principal investigator, perhaps over coffee, as soon as possible and ideally before starting work. The goal is to understand what specific duties and obligations are expected of each team member; how the lab works on a daily, weekly and monthly basis; and how the supervisor might handle typically sticky issues such as a non-renewed grant, long work hours or determining lead authorship of a paper. "If the postdoc or graduate student doesn't get anything from the PI about this, they need to ask for it," says Tony Nunez, a neuroscientist at Michigan State University (MSU) in East Lansing, where he also heads the postdoctoral office.

A similar fact-finding chat with a group of lab-mates can demonstrate how they function and interact. Louellen Essex, an executive coach in Minneapolis, Minnesota, says that the goal is to head off discord by learning the lie of the land before arguments have a chance to form. "The best way to deal with anything is before it happens," she says.

When joining a lab, junior researchers need to follow basic rules, says Wei-Chun Wei, a molecular cell biologist who is in her second postdoc at the University of Oxford, UK. "I always talk to an older lab member, usually the PI, to determine what those rules are," she says. "I ask about how the space is arranged, how to hook up the microscope, things like that."

But an initial meeting is not enough — the lines of communication need to stay open. Early-career researchers need to make sure that they are on track by checking in with their supervisor regularly, and not just in lab meetings that focus on research progress. Over time, the supervisor's requirements may shift. "You need to visit expectations periodically,"

says Nunez. "There is such a thing as memory decay, and things also change."

Keeping careful work records in a lab notebook or other medium can also help to stave off trouble. Jeremy Boss, an immunologist at



**"You want to understand where your boss is coming from."**

John Baldoni

Emory University in Atlanta, Georgia, recalls an awkward confrontation with a graduate student who seemed to be unproductive. "She started to tear up," he says. But then she pulled out her lab notebook and showed him detailed documentation of her work over the past year. "She was making great progress; it was just that I was unaware," says Boss.

Another way to avoid blow-ups is to keep a comprehensive shared calendar of the team's daily and weekly schedules and projects. Anna Kopeck, a pathology postdoc at MSU and co-chair of the university's postdoc association, says that her lab keeps such a list on a whiteboard. At a glance, everyone knows what everyone else — including the technicians — is doing, as well as when and for whom.

## SIMMER DOWN

So what happens when a conflict does arise with a colleague or supervisor? It is crucial to remain non-defensive and diplomatic, says Essex. "Don't come out confronting in a way that might be career-derailing," she says. It helps to calmly restate the other person's position (see 'How to handle friction'). "Paraphrase and ask questions if it's not clear, and

then respond with your point of view," she says. "The key is to be open to whatever it is they're saying."

If, for example, a principal investigator accuses a postdoc of failing to manage and mentor a graduate student effectively, the postdoc should stay calm, repeat the complaint and ask for clarification and examples. If the supervisor has the wrong impression or faulty information, the postdoc can clear up the misunderstanding.

But if the criticism is fair, the postdoc should agree and outline a remedy such as scheduling regular meetings with the student or sitting down to discuss his or her needs and presenting the outcome to the lab leader. The postdoc should make sure to thank the supervisor for bringing up the matter. "Staying off the defensive and being open to critical feedback is necessary to create productive professional relationships," says Essex. "No one is perfect."

When the junior researcher is the one broaching a sensitive topic or bringing up a complaint, Essex advises that he or she articulate the problem clearly and concisely, using conciliatory language such as, "I wonder if you might be open to," "Can you help me with" or "Do you think that", and presenting possible solutions. If, for example, the problem involves working extremely long hours, the junior researcher could suggest doing some of the work — data analysis, for example — from home. In some cases, requesting a shorter work week might be contrary to the lab culture; whether this is the case can often be discovered during the preliminary lab-group chat.

If the junior researcher is worn out and needs a few weeks off, it is best to arrange for lab-mates to pick up the slack during the absence before approaching the supervisor. "Tell the PI that you have it covered," says

## COMMUNICATION

### How to handle friction

As junior team members, graduate students and postdocs often face tricky moments with their principal investigators and lab-mates. Here are some ways to head off disputes and to manage them when they crop up.

**Set up an introductory chat.** Before the lab rotation or postdoc stint begins, arrange for a one-on-one meeting with your supervisor to clarify expectations and requirements such as working hours, performance assessments, what happens if a grant is not renewed and how lead authorship is decided.

**Meet lab-mates.** Get other members of the team together for a coffee before starting the post. Find out about the

principal investigator's foibles, quirks and hot buttons. How does he or she like the lab to be set up? What is the team dynamic like?

**Keep breathing.** Do not panic or get defensive if the supervisor or a lab-mate criticizes you. Take a breath and calmly paraphrase what was said. Try to avoid an argument by briefly explaining how you are handling the situation. If you were unaware of the problem, sketch out a plan of action and accept suggestions.

**Walk away.** If the other person is shouting, state that this is not the best time to discuss the issue and leave the room. A 10–15-minute stroll will help you to clear your head and think of solutions. **K.K.**





Wei-Chun Wei recommends learning about a lab's rules before starting work there.

Kopec. "You've talked with the other postdocs, the graduate students, the tech — and you can check on experiments from home and you have access to the data."

If an exchange gets heated, try stepping back, perhaps by telling the other person that this is not the right moment to discuss the subject, and suggesting another time to talk, Nunez counsels. Kopec says that when she is upset, she takes a deep breath and heads out for a walk. "Only deal with conflict when you're calm," she says. Use the interlude to contemplate solutions.

Publishing is a frequent source of conflict. Perhaps a postdoc or graduate student has been contacted by a journal about submitting a paper, but the supervisor objects because the journal is not high-profile enough. The exchange has the potential to escalate: the junior researcher needs the publication for his or her CV, but the principal investigator refuses to consider the journal, and one or both parties start to fume.

One option, says Kopec, is to consider doing more experiments, amassing stronger data and submitting to a better journal. But she adds that junior researchers need to get used to not always getting their own way: "You have to remember that your boss is still your boss at the end of the day."

It may also help to discuss any quarrel with a more senior lab member, administrator or mentor. "People need to know there's help for them when they're in conflict with a PI," says Kathy Barker, a lab-management specialist

in Seattle, Washington. "It's important to get out there, talk and get advice." The university postdoctoral office or international-scholar office (where applicable) can also provide support and advice.

At the very least, says Kopec, talking the problem through with a third party will give the postdoc a chance to vent and discuss it. The international-scholar office can help in administrative disputes involving visas, given that lab leaders often don't know the particulars of immigration law.

### CALLING IT QUITS

Sometimes, no matter how much an early-career researcher may have tried to soothe troubled waters, the relationship with his or her supervisor may be unworkable or irretrievably damaged. Changing labs is not the career-destroyer some might fear; staying in a conflict-ridden environment can be much worse (see *Nature* **442**, 324–325; 2006 and *Nature* **470**, 129–131; 2011). "If it's not working out in a lab, the postdoc and graduate student need to think about their career," says Boss. "It's a bad marriage and it needs to be annulled."

If quitting seems to be the best path, it is important to approach the issue tactfully and diplomatically, to give at least a few weeks' notice and to offer to finish a concrete set of tasks. "Don't say, 'Your lab is terrible, you're a monster and no one should work for you,'" says Boss.

Instead, he recommends that graduate students or postdocs explain that they have a certain number of years left before their lab rotation or postdoc is over, and that they would like to act now to make a transition into another area of research at another lab. "Defuse the conflict and take the high road," says Boss.

There may be logistical obstacles. If a postdoc is funded by a principal investigator's grant or being paid for by the university, the principal investigator should not object and there should be little problem in giving appropriate notice and switching to the new lab. But a postdoc fellowship might be restricted to research in a specific field or discipline, which would make it difficult to transfer to a lab doing different work. The fellowship's funding agency or foundation programme officer will be able to clarify the situation.

Ultimately, says Barker, managing conflicts is a finely nuanced balancing act. Early-career researchers must see to their own interests but avoid burning bridges with their principal investigators and other lab members. "Most bad situations will not fix themselves," she says. "You have to advocate for yourself." ■

**Karen Kaplan** is associate *Careers* editor at *Nature*.

### FACULTY

## Minority mentoring

A study finds that mentoring can help life-sciences faculty members at US institutions that serve ethnic minorities to win more grants, boost their publication rates and participate in more professional activities, including attending scientific conferences and developing collaborations (A. G. Campbell *et al.* *CBE Life Sci. Edu.* **12**, 394–402; 2013). The study, published on 4 September, partnered 32 faculty members from minority-serving institutions with established scientists at research-intensive universities for 8–10-week internships between 1997 and 2011. Participants' average number of publications rose from 0.84 before the internship to 1.37 afterwards. Their average number of grants increased from 0.06 to 0.59, up to four times the average of a control group of 129 faculty members at the same institutions who were not mentored.

### REPRODUCIBILITY

## Identification failure

Researchers are impeding reproducibility by not identifying experimental resources or properly sharing data, says a study (N. A. Vasilevsky *et al.* *PeerJ* **1**, e148; 2013). In the methods sections of 238 biomedical journal articles from 2012–13, 54% of resources such as antibody types and cell lines were not fully defined. A lack of identifiers also plagues online data-sharing, says lead author Nicole Vasilevsky, an ontologist at Oregon Health and Science University in Portland, adding that editors, funding agencies and grant reviewers should mandate resource identification. "If researchers don't keep track of what reagents they used, it could impact their own — and others' — research," she says.

### SALARY

## Maths degrees pay

US holders of master's degrees in maths earn more than those with degrees in biology or chemistry, finds a report, *Higher Education Pays*. Research firm College Measures in Rockville, Maryland, analysed first-year salaries for 15,118 master's holders in various fields from 2006 to 2011. Maths graduates typically earned US\$49,280–\$59,113, whereas biology graduates earned \$35,788–\$40,561 and chemistry graduates \$47,045–\$49,421. The numbers, says College Measures president Mark Schneider, imply high demand for mathematical problem-solving skills.



# REACH FOR THE STARS

*Impact factors.*

BY PRIYA CHAND

The holes widen into craters while I sleep.

The day is a blur — but that's school. No one pays attention anyway.

"No time for dinner," I say, my toes cold where they touch the linoleum.

Mother looks at me, wielding her spoon as the soup bubbles over. "I don't like this. Why can't you go natural?"

Unaugmented reflexes. The team would drop me. The schools wouldn't like that. Career suicide.

"I have to make captain."

She goes to eat and I go to the Chamber. Sit. Pincers lift my scalp away, layer by layer until the air hits my brain. I feel nothing. No pain.

Wires seep in, reinforcing my sensorimotor cortex. Brodmann's areas 3 and 4 — neurobiology was last year. Every detail is at the tip of my fingers, fresh as the day I learned it.

When the plate comes down to secure the electrodes, my foot twitches. The device locks, then my skull is replaced. Back to normal.

No, not normal. The only parts of the Moon we've explored are its impact craters.

No one can beat me. We're all augmented, of course, but they haven't spent hours brainmapping and it shows. One girl keeps driving the ball into the ground. Should've elevated precision. Idiot.

Later, we have pizza to celebrate. I can't remember how to eat it, but I watch the others and my mirror neurons fire zip-zip-zip and it's there. Fold and cram.

Mother is making tea when I get home. I tell her we won. She nods and sniffs the pot, adding spices I've never had the time to learn about. Still, the scents are soothing. I yawn.

The stimulator disintegrates during sleep. There'll be a fine metal dust on my bed in the morning. Mother changes the sheets every day, I watch her tuck the linen in while I brush my teeth and comb my hair.

I wake to a pale arm near my body. Shriek — it's creepy. Mother runs in, brays nervous laughter and as my synapses reconnect I start laughing too.

It's mine. My arm.

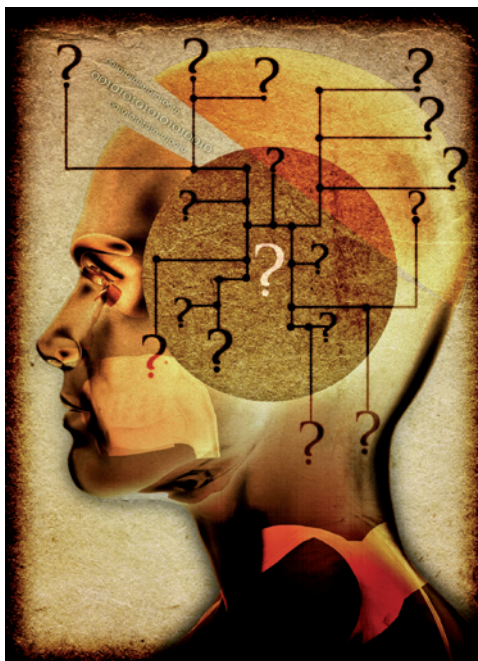
Test Wednesday.

Active augmentation

is banned so I use it

to study, to help me

best comprehend maths and higher logic. The stimulators — two this time — target my thalamus, strengthening the cortico-thalamic loops. This is more effective than direct stimulation of the prefrontal cortex. There are plenty of data to confirm.



As the machine disengages, I hear a slickness, a 'gloop' where the air rushes to fill the gaps between torn brain cells.

Everyone does it. It's perfectly normal.

Studying takes 20 minutes. The memories slide into vacant spots, the barrage of signaling drenches my hippocampus. I feel those pathways strengthen, standing out like veins on a bodybuilder. My mind is a muscle.

I ace the test. I come home to tell the woman in the kitchen about it. She says I call her Mother.

"Like mater," I say. Latin last semester, an accelerated programme for premeds. They tell us to come to class augmented, but you can't practise medicine on stimulators. I'm waiting for the FDA to finish phase IV trials. It'll be legal soon.

"Yes, exactly," she says. Mother. Mater.

Dura and pia, my hard and soft mothers; I wonder where she fits in, her sharp bones and unexpected curves. I look for her in the layers of my head, which fold back for the anatomy pre-practical.

Fine motor skills.

My hands don't shake. I lay out the fetal

pig's heart, ventricles exposed. I extract its smooth brain, a strange grey-yellow in a formalin jug.

The next morning, I can't tell left from right.

"There's a special report on," Mother/Mater tells me over breakfast. I take the cereal one bite at a time. Yes, I hold spoons with my right hand. Problem solved. "The postmortem on John Li."

John Li wanted to get into Harvard. John Li got addicted to augmentation, skipping sleep so the stimulators wouldn't disintegrate overnight.

And then he killed himself.

"I don't have the time."

"It's the weekend."

"Finals are next week." One slot, one thousand students. I have to be the most perfect of all.

Huddled over my notes on the renal system, drops of blood where the sealant missed gluing my skin together, I hear the TV. Expert commentary on John Li's brain. "Scarred."

I imagine the Moon, welcoming every meteor to its blank surface. No strata for protection. You can't be a blank slate forever.

Last day of exams.

I wake to metallic dust on my pillow. I sleep on clean sheets; we must have a robot for that.

Brush, eat, test.

The second it ends, I follow 30 others to the bathroom. We discuss how poorly scanned the holograms were, blurring the lens, stroma and cornea together, ha-ha isn't that ironic.

The joke falls flat. There's too much awareness of how little separates us. We can't wish each other good luck; that's the only difference between getting in and not getting in when you're identically gifted.

The proctors shoot disapproval as we wipe our ears, but this is the future. Silver smudges on our fingers and brains like the Moon.

Some want to ban nontherapeutic stimulation, but that's pointless. Normal students forget plenty, no studies show significant deviation from the norm.

We are the future.

I look in the mirror to straighten my hair, but I can't tell which anxious face is mine. ■

**Priya Chand** graduated from the University of Chicago with a degree in neuroscience. She's since found that writing stories can be as mindbending as conducting experiments, and is looking for a way to combine the two. Find her @writelies on Twitter.

JACEY

➔ NATURE.COM

Follow Futures:

@NatureFutures

go.nature.com/mtoodm